

EP 5566525 (1)
C12N15/00F12824-C12N15/82-
[MC7K215:00]-[MC7K207:00]-
[MC7K213:00]-



C12N15/82

(11) Publication number: 0 566 525 A2

- 2- *-

EUROPEAN PATENT APPLICATION

(12)

(21) Application number: 93810190.4

(51) Int. Cl.⁵: C12N 15/40, C12N 15/82,
C12Q 1/70, A01H 5/00

(22) Date of filing: 16.03.93

The applicant has subsequently filed a
sequence listing and declared, that it includes
no new matter.

(30) Priority: 19.03.92 GB 9206016

(43) Date of publication of application:
20.10.93 Bulletin 93/42

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE IT LI LU NL
PT SE

(71) Applicant: SANDOZ LTD.
Lichtstrasse 35
CH-4002 Basel (CH)
(84) BE CH DK ES FR GB GR IE IT LI LU NL PT SE

(71) Applicant: SANDOZ-PATENT-GMBH
Humboldtstrasse 3
D-79539 Lörrach (DE)
(84) DE

(71) Applicant: SANDOZ-ERFINDUNGEN
Verwaltungsgesellschaft m.b.H.
Brunner Strasse 59
A-1230 Wien (AT)
(84) AT

(72) Inventor: Van Grinsven, Martinus Quirinus
Joseph Marie
Wezenland 5
NL-1602 MA Enkhuizen (NL)
Inventor: De Haan, Petrus Theodorus
Kruideel 34
NL-1602 GL Enkhuizen (NL)
Inventor: Gielen, Johannes Jacobus
Ludgerus
Jan Gooskaai 73
NL-1602 GC Enkhuizen (NL)
Inventor: Peters, Dirk
Edelmanlaan 4
NL-6703 EX Wageningen (NL)
Inventor: Goldbach, Robert Willen
Hollandseweg 159
NL-4705 BC Wageningen (NL)

(54) Recombinant tospovirus DNA constructs and plants comprising such constructs.

(57) Recombinant Impatiens Necrotic Spot Virus (INSV) DNA constructs comprising an INSV DNA coding for transcription into INSV RNA sequences or into RNA sequences related thereto, the use of such DNA constructs to transform plants having reduced susceptibility to INSV infection and probes for the isolation of INSV or diagnosis of plant INSV related diseases.

EP 0 566 525 A2

The present invention relates to plants having reduced susceptibility to infection from tospoviruses, genetic material capable of generating tolerance to tospoviruses, probes suitable for isolating and diagnosing, and processes for obtaining such plants and genetic material and probes.

Viral infections in plants are frequently responsible for detrimental effects in growth, undesirable morphological changes, decreased yield and the like. Such infections often result in a higher susceptibility to infection in infected plants to other plant pathogens and plant pests. Transmission of plant viruses generally occurs via insect or fungal carriers or may occur through mechanical means.

Plant breeders continuously look to develop varieties of crop plant species tolerant to or resistant to specific virus strains. In the past, virus resistance conferring genes have been transferred from wild types related to commercial plants into commercial varieties through breeding. The transfer of an existing resistance in the wild from the wild type gene pool to a cultivar is a tedious process in which the resistance conferring gene(s) must first be identified in a source (donor) plant species and then combined into the gene pool of a commercial variety. Resistance or tolerance generated in this way is typically active only against one or at best a few strains of the virus in question. One disadvantage of breeding cultivars for resistance to a particular virus species is that there is often a lack of a gene source suitable for conferring disease resistance within the crop species.

Other approaches to limit the effect of virus induced disease on plants include the use of chemicals such as insecticides, fungicides and the like which act against virus carriers, and/or rely on the employment of preventative methods such as efficient phytosanitary working conditions. However, the use of chemicals to combat virus disease by killing the carrier is subject to increasingly tougher governmental regulations which present growers with a decreasing scale of permitted chemical plant-protectants.

In an alternative, a system referred to as "cross-protection" may be employed. Cross-protection is a phenomenon in which infection of a plant with one strain of a virus protects that plant against superinfection with a second related virus strain. The cross-protection method preferentially involves the use of avirulent virus strains to infect plants, which act to inhibit a secondary infection with a virulent strain of the same virus. However, the use of a natural cross-protection system can have several disadvantages. The method is very labour intensive because it requires inoculation of every plant crop, and carries the risk that an avirulent strain may mutate to a virulent strain, thus becoming a causal agent for crop disease in itself. A further possible hazard is that an avirulent virus strain in one plant species can act as a virulent strain in another plant species.

Several studies have indicated that the viral coat protein of the protecting virus plays an important role in cross-protection and that protection occurs when the resident virus and the challenging virus have the same or closely related coat protein structures.

Recent developments in gene manipulation and plant transformation techniques have given rise to new methods for generating virus resistance in plants. Genetically engineered cross-protection is a form of virus resistance which phenotypically resembles natural cross-protection, but is achieved through the expression of genetic information of a viral coat protein from the genome of a genetically manipulated plant. Generation of virus resistance via genetic engineering has been described in for instance, EP 223 452 and reported by Abel et al [(1986) Science 232:738-743]. It was shown that expression of the tobacco mosaic virus strain U1 (TMV-U1) coat protein gene from the genome of a transgenic plant resulted in a delay of symptom development after infection with any TMV strain. Similar results with respect to coat protein-mediated protection have also been obtained for alfalfa mosaic virus (AMV), potato virus X (PVX) and cucumber mosaic virus (CMV).

Although TMV, CMV, AMV and PVX belong to different virus groups, they share a common architecture: in all such viruses the viral RNA is a positive strand RNA encapsidated by a viral coat consisting of many individual but identical viral coat proteins.

However, tospoviruses are essentially different from the plant viruses mentioned above. The genus tospovirus belongs to the family Bunyaviridae. All tospoviruses are transmitted by thrips. The virus particles are spherical in shape (80-120 nm in diameter) and contain internal nucleocapsids surrounded by a lipid envelope studded with glycoprotein surface projections. The multipartite genome consists of linear single stranded RNA molecules of negative or ambisense polarity. The terminal nucleotides of these RNA molecules are characterised by a consensus sequence as follows: 5' AGAGCAAUX.....GAUUGCUCU 3', wherein X is C or U. Members of the tospovirus group include tomato spotted wilt virus (TSWV), Impatiens necrotic spot virus (INSV), and tomato chlorotic spot virus (TCSV), also known as tomato mottled spot virus (TMSV) or TSWV-like isolate BR-O3. A general description of a tospovirus, using TSWV as a representative of the genus tospoviruses can be found in our co-pending application EP 426 195 herein incorporated by reference.

The tospovirus particle contains at least 4 distinct structural proteins: an internal nucleocapsid protein N of 29 kd and two membrane glycoproteins: G1, approximately 78 kd, and G2 approximately 58 kd. In addition, minor amounts of a large protein, L, approximately 260 kd have been detected in virus particles. Tospoviral genomes consist of three linear single stranded RNA molecules of about 2900 nucleotides (nt) (S RNA), about 5000 nt, (M RNA) and about 8900 nt (L RNA), each tightly associated with nucleocapsid proteins and a few

copies of the L protein to form circular nucleocapsids. A schematic structure outlining most properties of an INSV is given in Figure 1. Based on the above and other properties, INSV (like TSWV) has been classified as a member of the tospovirus genus.

5 Circumstantial evidence has been presented which suggests that an M RNA encoded gene is directly or indirectly involved in the synthesis of the G1 membrane glycoprotein [Verkleij and Peters, (1983) J. Gen. Virol. 64:677-686].

As mentioned above, tospoviruses such as TSWV, INSV and the like are transmitted by certain species of thrips. These tospovirus carriers belong to the family *Tripidae* and include tobacco thrips (*Frankliniella fusca* (Hinds.)), western flower thrips (*F. occidentalis* (Pergande)), common blossom thrips (*F. Schultzei* (Trybom)), chilli thrips (*Scirtothrips dorsalis* (Hood)), *Thrips setosus* (Moulton), onion thrips (*T. tabaci* (Lindeman)), *F. intonsa* and melon thrips (*T. palmi* (Karny)). The tospovirus is acquired by thrips only during their larval stages. Larvae can transmit the virus before they pupate but adults more commonly transmit the virus. Adult thrips can remain infective throughout their lives.

15 Tospoviruses are widespread in temperate, subtropical and tropical climate zones throughout the world. The current distribution of tospoviruses covers all continents and makes them one of the most widely distributed of groups of plant viruses. At least 370 plant species representing 50 plant families, both monocotyledons and dicotyledons, are naturally infected by tospoviruses of the *Bunyaviridae*. Tospoviruses seriously affect the production of food and ornamental crops. Symptoms of tospovirus infection in plants include stunting, ring-
20 spots, dark purple-brown sunken spots, stem browning, flower breaking, necrotic and pigmental lesions and patterns, yellows and non-necrotic mottle, mosaic in greens or even total plant death. Most plant hosts display only a few of these symptoms, however, the wide range of symptoms produced by tospovirus infection has complicated diagnosis of the disease and has led to individual diseases being given several different names. A further complication is that tospovirus symptoms within the same plant species may vary depending on the
25 age of the plant, time of infection during the life-cycle of the plant, nutritional levels, environmental conditions, such as temperature, and the like.

Although TSWV has been known for many years, is widely distributed, and is the causal agent of a disease which leads to significant loss in yield in crops and ornamentals, limited progress has been made in identifying sources of genes capable of conferring resistance to TSWV or other tospoviruses. A monogenic TSWV tolerance has been identified in *Lycopersicon peruvianum*, but this trait has not been transferred to cultivated tomatoes so far, nor has a resistance source been identified for other crop species. The use of natural cross-protection systems to decrease the invasive effects by tospovirus strains capable of causing damage is not well documented. Limited positive results have been reported for tomato and lettuce.

35 The introduction of genetic information capable of conferring resistance or tolerance to tospoviruses into plant gene pools by means of genetic manipulation provides the breeder and grower alike with a new method for combatting tospovirus induced disease. In particular, it has been found that genetic manipulation techniques may be employed to confer resistance to INSV related disease in plants.

Detailed Description

40 According to the present invention there is provided a recombinant INSV DNA construct comprising a DNA sequence coding for transcription into

- a) an RNA sequence of an INSV or an RNA sequence homologous thereto;
- b) an RNA sequence of an INSV or an RNA sequence homologous thereto capable of encoding for an INSV protein or a part thereof, in which one or more codons have been replaced by synonyms, or an RNA sequence homologous thereto; or
- 45 c) an RNA sequence complementary to an RNA sequence according to a) or b),

which INSV DNA is under expression control of a promoter capable of functioning in plants and includes a terminator capable of functioning in plants.

50 The DNA sequences defined under a), b) and c) above, for the purposes of the present invention will be referred to as "INSV Related DNA Sequences" hereinafter. An INSV Related DNA Sequence according to the invention may be modified as appropriate to create mutants or modified sequences homologous to such INSV Related DNA Sequences from which they are derived, using methods known to those skilled in the art such as site-directed mutagenesis and the like. Such mutants or modified coding sequences are embraced within
55 the spirit and scope of the invention.

The term "RNA sequence of an INSV" may refer to a sequence of the S, M or L RNA strand, preferably an S or M RNA strand, more preferably to an S RNA strand of an INSV.

The term "RNA sequence homologous to an RNA sequence of an INSV" refers to an RNA sequence of an INSV wherein a number of nucleotides have been deleted and/or added but which is still capable of hybrid-

ization to a nucleotide sequence complementary to an RNA sequence of an INSV under appropriate hybridization conditions. For the purposes of the present invention appropriate hybridization conditions may include but are not limited to, for example, an incubation for about 16 hours at 42°C, in a buffer system comprising 5 x standard saline citrate (SSC), 0.5% sodium dodecylsulphate (SDS), 5 x Denhardt's solution, 50% formamide and 100 µg/ml carrier DNA (hereinafter the buffer system), followed by washing 3x in buffer comprising 1 x SSC and 0.1% SDS at 65°C for approximately an hour each time.

Preferably, hybridization conditions employed in the present invention may involve incubation in a buffer system for about 16 hours at 49°C and washing 3x in a buffer comprising 0.1 x SSC and 0.1% SDS at 55°C for about an hour each time. More preferably, hybridization conditions may involve incubation in a buffer system for about 16 hours at 55°C and washing 3x in a buffer comprising 0.1 x SSC and 0.1% SDS at 65°C for approximately an hour each time.

The length of the INSV Related DNA Sequence will i.a. depend on the particular strategy to be followed, as will become apparent from the description hereinafter. In general, the INSV Related DNA Sequence may comprise at least 20, and suitably 50 or more nucleotides.

The term "promoter" refers to the nucleotide sequence upstream from the transcriptional start site and which contains all the regulatory regions required for transcription, including the region coding for the leader sequence of mRNA (which leader sequence comprises the ribosomal binding site and initiates translation at the AUG start codon).

Examples of promoters suitable for use in DNA constructs of the present invention include viral, fungal, bacterial, animal and plant derived promoters capable of functioning in plant cells. The promoter may express the DNA constitutively or differentially. Suitable examples of promoters differentially regulating DNA expression are promoters inducible by disease carriers, such as thrips, e.g. so-called wound-inducible promoters. It will be appreciated that the promoter employed should give rise to the expression of an INSV Related DNA Sequence at a rate sufficient to produce the amount of RNA necessary to decrease INSV susceptibility in a transformed plant. The required amount of RNA to be transcribed may vary with the type of plant. Particularly preferred promoters include the cauliflower mosaic virus 35S (CaMV 35S) promoter, derivatives thereof, and a promoter inducible after wounding by a disease carrier such as thrips, e.g. a wound inducible promoter. Examples of further suitable promoters include nopaline synthase, octopine synthase and the like.

The term "terminator" refers to a DNA sequence at the end of a transcriptional unit which signals termination of transcription. Terminators are DNA 3'-non-translated sequences that contain a polyadenylation signal, that causes the addition of polyadenylate sequences to the 3'-end of a primary transcript. Terminators active in plant cells are known and described in the literature. They may be isolated from bacteria, fungi, viruses, animals and/or plants. Examples of terminators particularly suitable for use in the DNA constructs of the invention include the nopaline synthase terminator of *A. tumefaciens*, the 35S terminator of CaMV and the zein terminator from *Zea mays*.

In accordance with the present invention, an RNA sequence is complementary to another RNA sequence if it is able to form a hydrogen-bonded complex therewith, according to rules of base pairing under appropriate hybridization conditions (as described hereinabove).

The present invention also provides a vector capable of introducing the DNA construct of the invention into plants and methods of producing such vectors.

The term "vector" as employed herein refers to a vehicle with which DNA constructs of INSV or fragments thereof may be incorporated into the cells of a host organism.

The term "plants" refers to differentiated plants as well as undifferentiated plant material such as protoplasts, plant cells, including cybrids and hybrids, seeds, plantlets and the like which under appropriate conditions can develop into mature plants, progeny thereof and parts thereof such as cuttings, fruits of such plants and the like.

The invention further provides plants comprising in their genome a DNA construct of the invention, and methods of producing such plants. Such methods include plant breeding, plantlets derived from protoplast fusion and the like.

The plants according to the invention have reduced susceptibility to diseases induced by INSV or diseases related to INSV infection and suffer from substantially fewer or none of the disadvantages and limitations of plants obtained by classical methods as mentioned hereinabove.

Many types of plants are susceptible to INSV infection however only in some types is INSV infection known to give rise to a disease state directly attributable to the virus.

Such types of plants include the ornamental or flowering plants. Examples of such plants include but are not limited to Ageratum, Amaranthus, Anthirrhinum, Aquilegia, Begonia, Chrysanthemum, Cineraria, clover, Cosmos, cowpea, Cyclamen, Dahlia, Datura, Delphinium, Gerbera, Gladiolus, Gloxinia, Hippeastrum, Impatiens, Mesembryanthemum, petunia, Primula, Saint Paulia, Salpiglossis, Tagetes, Verbena,

Viola, Vinca, Zinnia, Pelargonium and the like.

Other types of plants may be susceptible to INSV infection but these plants may not present disease symptoms directly associated with INSV infection, however such plants may present symptoms of a disease as a result of a secondary infection by a different organism made possible as a result of an initial infection by INSV. Such plants may therefore be viewed as being the subject of an INSV infection related disease and may include plants selected from a wider group of plant types. Further examples of this group of plant types may include vegetable and other crops. Such crop types include alfalfa, aubergine, beet, broad bean, broccoli, brussels sprouts, cabbage, cauliflower, celery, chicory, cow pea, cucumber, endive, gourd, groundnut, lettuce, melon, onion, papaya, pea, peanut, pepper, pineapple, potato, safflower, snap bean, soybean, spinach, squash, sugarbeet, sunflower, tobacco, tomato, water melon and the like.

The invention relates in particular to ornamental plants and preferably to those listed ornamental plants comprising in their plant genome a DNA construct of the invention.

The particular features of tospoviruses including those of INSV are illustrated hereinafter.

The S, M and L RNA are single stranded RNA molecules. The S RNA of INSV is about 3000 nucleotides long (SEQ ID No.1; SEQ ID No. 2) and comprises two genes, one (SEQ ID No.3) encoding a non-structural protein (NSs) in viral sense, the other one (SEQ ID No.11) encoding the nucleocapsid protein (N) in viral complementary sense. The intergenic region between the NSs- and N-gene can be folded into a secondary structure (Seq ID No. 7 and SEQ ID No.8). The 5'- and 3'-terminal sequences of the S RNA are capable of hybridizing to each other such that the first nucleotide is opposite (and complementary) to the last nucleotide of said S RNA strand. For the purposes of the description the double-stranded structure obtained by hybridizing both RNA termini will be referred to as a "pan-handle" (SEQ ID No.5 and SEQ ID NO. 6) hereinafter.

The M RNA strand of INSV comprises about 5000 nucleotides (SEQ ID No. 14). It contains at least two open reading frames, one encoding a non-structural protein (NSm) in viral sense (SEQ ID No.15), and another open reading frame (SEQ ID No.21) in viral complementary sense. This open reading frame is translated on polysomes located on the endoplasmic reticulum where the nascent polypeptide chain is cleaved co-translationally to form the spike proteins G1 and G2 respectively. As with S RNA, the termini of the M RNA strand are complementary to each other and may likewise hybridize to form a "pan-handle" (SEQ ID No.18 and SEQ ID No.19).

The L RNA strand of INSV comprises about 8900 nucleotides. It contains complementary 3' and 5' ends for a length of from about 50 to about 80 nucleotides. The RNA has a negative polarity, with one open reading frame (ORF) located as the viral complementary strand. This ORF corresponds to a primary translation product of about 2875 amino acids in length with an anticipated Mw of between about 300,000 to about 350,000. Comparison with the polymerase proteins of other negative strand viruses indicates that this protein probably represents a viral polymerase. In some mutant strains, shortened L RNA molecules have been found in addition to the wild type, full length L RNA. These shortened L RNAs however are observed to possess the characteristic terminal nucleotide sequences and thus are capable of forming "pan handle" structures. They are also encapsidated with nucleocapsid protein and are included in virus particles. Their presence suppresses symptom development resulting in less severe detrimental effect. Thus, these shortened L RNA molecules can be regarded as defective interfering (DI) RNAs. A defective interfering RNA is one which is capable of interfering in replication by competing with other genomic RNAs for polymerases and therefore is capable of being replicated, and by so doing inhibits the replication and/or expression of other genomic RNAs with which it is competing. Thus, a DI RNA may comprise any RNA sequence which is capable of being replicated and may be an L, S, or M RNA within the context of the present invention. Such DI RNA sequences may comprise RNA sequences which have had nucleotides either deleted from or added thereto provided that they are capable of competing for polymerases and of replicating.

A preferred embodiment of the invention relates to DNA constructs of the invention coding for transcription into INSV RNA sequences of a "pan-handle" (SEQ ID No.5, SEQ ID No.6; SEQ ID No.18, SEQ ID No.19), or into INSV RNA sequences homologous thereto.

Another preferred embodiment of the invention relates to DNA constructs of the invention coding for transcription into INSV-RNA sequences of an open reading frame in viral complementary sense i.e. having negative polarity, or into corresponding RNA sequences in which one or more codons have been replaced by their synonyms, or into RNA sequences homologous thereto.

A further preferred embodiment of the invention relates to DNA constructs of the invention coding for transcription into INSV-RNA sequences of a hairpin (SEQ ID No.7, SEQ ID No.8; SEQ ID No.13, SEQ ID No.16), or into RNA sequences homologous thereto.

Preferably, the INSV-RNA sequence referred to hereinabove has at least 20 nucleotides. Preferably, the INSV-RNA sequence has at least 50 nucleotides.

Examples of DNA constructs suitable for use according to the invention include INSV-Related DNA Se-

quences coding for transcription into (reference is made to the sequence listing):

- i) the viral S RNA nucleotide sequence from 1 to 3017 (SEQ. ID No.1)
- ii) the viral S RNA nucleotide sequence from position 25 to 3017 (SEQ. ID No.2);
- iii) the viral S RNA nucleotide sequence from 87 to 1436 (SEQ. ID No.3);
- iv) the viral S RNA nucleotide sequence from 2080 to 2868 (SEQ. ID No.4);
- v) the viral S RNA "pan-handle" structure comprising:
 - a) a first nucleotide sequence of from about 30 to about 36 nucleotides in length from the 5' end of the viral S RNA
 - and
 - b) a second nucleotide sequence of from about 30 to about 36 nucleotides in length from the 3' end of the viral S RNA
- vi) the viral S RNA nucleotide sequence from 1437 to 2079; (SEQ ID No. 7)
- vii) the viral S RNA nucleotide sequence from 1440 to 2041; (SEQ ID No.8)
- viii) the viral complementary S RNA nucleotide sequence from 1 to about 3017; (SEQ ID No.9)
- ix) the viral complementary S RNA nucleotide sequence from 1 to 2993; (SEQ ID No.10)
- x) the viral complementary S RNA nucleotide sequence from 150 to 938; (SEQ ID No.11)
- xi) the S RNA nucleotide sequence from 1581 to 2930 of the viral complementary S RNA strand; (SEQ ID No.12);
- xii) the viral complementary S RNA secondary structure having a nucleotide sequence of 642 nucleotides from 939 to 1580; (SEQ ID No.13)
- xiii) S RNA nucleotide sequence from 87 to 1436 in which one or more codons have been replaced by their synonyms;
- xiv) S RNA nucleotide sequence from 2080 to 2868 in which one or more codons have been replaced by their synonyms;
- xv) the M RNA nucleotide sequence from 1 to 4970 (SEQ ID No.14);
- xvi) the M RNA sequence from 86 to 997 (SEQ ID No.15);
- xvii) the M RNA sequence of the intergenic region from 998 to 1470 (SEQ ID No.16);
- xviii) the M RNA sequence from 1471 to 4884; (SEQ ID No. 17)
- xix) the M RNA "pan-handle" structure comprising : a) a first nucleotide sequence of from about 30 to about 36 nucleotides in length from the 5' end of the viral M RNA
- and
- b) a second nucleotide sequence of from about 30 to about 36 nucleotides in length from the 3' end of the viral M RNA
- xx) the complementary viral M RNA sequence from 1 to 4970; (SEQ ID No.20)
- xxi) the complementary viral M RNA sequence from position 87 to position 3500 of the complementary viral M RNA sequence; (SEQ ID No.21)
- xxii) the complementary viral M RNA sequence from position 3974 to 4885 (SEQ ID No.22)
- xxiii) RNA sequences homologous to the nucleotide sequences defined under i) to xii) and xv) to xxii) hereinabove.
- xxiv) fragments of sequences defined under i) to xxii) hereinabove.

Preferred INSV-Related DNA Sequences code for transcription into the RNA sequences according to sequences iv) to xii) and xv) to xxii) as defined above, or into RNA sequences homologous thereto, or into fragments thereof comprising at least 15 nucleotides, more preferably at least 20 nucleotides, and most preferably at least 50 nucleotides.

According to another preferred embodiment of the invention the DNA constructs of the invention comprise INSV Related DNA Sequences coding for transcription into a combination of the 5' and 3' terminal sequences (ie "pan-handles") of viral S, M or L RNA respectively, more preferably of S or M RNA, and most preferably of S RNA . Examples of S RNA and M RNA terminal sequences include

- i) a first nucleotide sequence 36 nucleotides in length from the 5' end of the viral S RNA :

5' AGAGCAATNN NNNNNNNNNN NNNNGAACAAAC CCAAGC 3'

(SEQ ID No.5 ie nucleotides from position 1 to 36 of SEQ ID No.1 , where N stands for A,T,G,or C)

and

a second nucleotide sequence 36 nucleotides in length from the 3' end of the viral S RNA:

5' GATTATATG ATGTTATATT CGTGACACAA TTGCTCT 3'

- (SEQ ID No.6 ie nucleotides from position 2981 to 3017 of SEQ ID No.1)
 ii) a first nucleotide sequence of 36 nucleotides in length from the 5' end of the viral M RNA :

5' AGAGCAATCA GTGCATCAA ATTATATCTA GCCGAA 3'

(SEQ ID No.18 ie nucleotides from position 1 to 36 of SEQ ID No.13)

and

- b) a second nucleotide sequence 36 nucleotides in length from the 3' end of the viral M RNA

5' TGTGTATGT AGAGATTTTG TTTGCACTGA TTGCTC T 3'

(SEQ ID No.19 ie nucleotides from position 4941 to 4970 of SEQ ID No. 13)

In the case of the terminus at the 5' end of the S RNA it is not known whether or not there are sixteen or seventeen nucleotides in the unknown region demarked by a series of "N"s, however the exact number of nucleotides in this region is not considered to be critical to the formation of "pan-handle" structures so long as the 5' end of the S RNA is capable of complementing the 3' end of the S RNA thus enabling the formation of a "pan-handle" structure.

The invention further provides probes suitable for use as diagnostic tools for the diagnosis of disease in plants suspected of being infected with INSV tospoviruses. Such probes comprise a labeled oligonucleotide (RNA or DNA) sequence complementary to an RNA sequence of an INSV tospovirus. The desired length of the sequence and appropriate method for diagnostic use of probes are known by those skilled in the art. A suitable probe may comprise a nucleotide sequence of at least 12 to about 800 nucleotides, preferably at least 15, more preferably more than 30 nucleotides, and most preferably from about 400 to 600 nucleotides complementary to an RNA sequence of an INSV tospovirus.

Probes according to the invention are helpful in identifying INSV tospovirus RNA or parts thereof in infected plant material i.a. for diagnostic purposes prior to full presentation of disease symptoms in plants.

The invention accordingly also provides a diagnostic method of determining INSV tospovirus infection in plants which comprises detecting INSV tospovirus replicative forms employing the probes of the invention in dot-blot type assays.

Probes according to the invention are useful in the construction of and use of chimeric genes comprising a DNA sequence corresponding to an RNA sequence of an INSV tospovirus.

The DNA constructs of the invention may be obtained by insertion of an INSV Related DNA Sequence in an appropriate expression vector, such that the sequence is brought under expression control of a promoter capable of functioning in plants and its transcription is terminated by a terminator capable of functioning in plants.

The term "appropriate expression vector" as used herein refers to a vector containing a promoter region and a terminator region which are capable of functioning in plant cells.

The insertion of an INSV Related DNA Sequence into an appropriate expression vector may be carried out in a manner known per se. Suitable procedures are illustrated in the examples hereinafter.

Likewise the construction of an appropriate expression vector may be carried out in a manner known per se.

Plants according to the invention may be obtained by

- a) inserting into the genome of a plant cell a DNA construct as hereinbefore defined;
- b) obtaining transformed cells; and
- c) regenerating from the transformed cells genetically transformed plants.

DNA vectors of the present invention may be inserted into the plant genome of plants susceptible to INSV infection. Such plant transformation may be carried out employing techniques known per se for the transformation of plants, such as plant transformation techniques involving Ti plasmids derived from *Agrobacterium tumefaciens*, *A. rhizogenes* or modifications thereof, naked DNA transformation or electroporation of isolated plant cells or organized plant structures, the use of micro-projectiles to deliver DNA, the use of laser systems, liposomes, or viruses or pollen as transformation vectors and the like.

Plants of the invention may be monitored for expression of an INSV-Related DNA Sequence by methods known in the art, including Northern analysis, Southern analysis, PCR techniques and/or immunological techniques and the like. The plants of the invention show decreased susceptibility to INSV infection as demonstrated by tests whereby the plants are exposed to INSV preferentially at a concentration in the range at which the rate of disease symptoms correlates linearly with INSV concentration in the inoculum.

Methods suitable for INSV inoculation are known in the art and include mechanical inoculation, and in particular, the use of appropriate vectors.

Plants of the invention may also be obtained by the crossing of a plant obtained according to the methods of the invention with another plant to produce plants having in their plant genome a DNA construct of the invention.

5 The invention is illustrated by the following non-limiting examples and accompanying figures.

Figure 1: Schematic representation of an INSV particle .

Figure 2: Sequence strategy for INSV viral S RNA.

Figure 3: Open reading frame analysis of the INSV S RNA, full bars represent translational stop codons (TAA, TAG, TGA), half size bars indicate start codons (ATG).

10 Figure 4: Schematic review of the construction of a suitable expression vector (pZU-B).

Figure 5: Schematic review of the construction of a suitable plasmid comprising the INSV N protein-coding sequence.

Figure 6: Schematic review of the construction of a suitable plasmid comprising the INSV NSs protein-coding sequence.

15 Figure 7: Schematic review of the construction of a suitable plasmid comprising the INSV NSm protein-coding sequence.

Figure 8: Schematic review of the construction of a suitable plasmid comprising the INSV G1/G2 glycoprotein precursor-coding sequence.

Figure 9: Schematic review of the construction of a INSV N gene-containing plant transformation vector.

20 Figure 10: Schematic review of the construction of a INSV NSs gene-containing plant transformation vector.

Figure 11: Schematic review of the construction of a INSV G1/G2 glycoprotein precursor gene-containing plant transformation vector.

25 Figure 12: Schematic review of the construction of a INSV NSm gene-containing plant transformation vector.

Figure 13: The secondary structure located at the intergenic region of INSV S RNA.

30 Suitable examples of preferred INSV Related DNA Sequences coding for transcription into a sequence of the secondary structure of the intergenic region of S RNA or of RNA sequences homologous thereto are sequences coding for the 1437 to 2079 nucleotide sequence of S RNA or for a sequence homologous to such sequences.

Other advantageous features of the present invention will be apparent from the following examples.

MATERIAL AND METHODS

35 All INSV RNA-derived sequences presented here are depicted as DNA sequences for the sole purpose of uniformity. It will be appreciated that this is done for convenience.

Cultivars of *Nicotiana tabacum* and *Petunia hybrida*, used in plant transformation studies, are grown under standard greenhouse conditions. Axenic explant material is grown on standard MS media [Murashige and

40 *Skoog*, (1962) *Physiol Plant* 15: 473-497] containing appropriate phytohormones and sucrose concentrations. *E. coli* bacteria are grown on rotary shakers at 37°C in standard LB-medium. *Agrobacterium tumefaciens* strains are grown at 28°C in MinA medium supplemented with 0.1 % glucose [Ausubel et al., (1987) *Current Protocols in Molecular Biology*, Green Publishing Associates and Wiley Intersciences, New York, Chichester, Brisbane, Toronto, and Singapore].

45 In all cloning procedures the *E. coli* strain JM83, (F⁻, Δ(lac-pro), ara, rpsL, Ø80, dlacZM15) is used as a recipient for recombinant plasmids.

Binary vectors are conjugated to *Agrobacterium tumefaciens* strain LBA 4404, a strain containing the Ti-plasmid vir region, [Hoekema et al., (1983) *Nature* 303:179-180] in standard triparental matings using the *E. coli* HB101, containing the plasmid pRK2013 as a helper strain. [Figurski and Helinski, (1979) *Proc. Natl. Acad. Sci. USA* 76:1648-1652] Appropriate *Agrobacterium tumefaciens* recipients are selected on media containing rifampicin (50 µg/ml) and kanamycin (50 µg/ml).

50 Cloning of fragments in the vectors pUC19 [Yanish-Perron et al. (1985) *Gene* 33:103-119], pBluescript (Stratagene), pBIN19 [Bevan et al., (1984) *Nucl Acids Res.* 12:8711-8721] or derivatives, restriction enzyme analysis of DNA, transformation to *E. coli* recipient strains, isolation of plasmid DNA on small as well as large scale, nick-translation, *in vitro* transcription, DNA sequencing, Southern blotting and DNA gel electrophoresis are performed according to standard procedures [Maniatis et al., (1982) *Molecular Cloning*, a Laboratory Manual, Cold Spring Harbor Laboratory, New York; Ausubel et al. supra, (1987)].

55 DNA amplification using the polymerase chain reaction (PCR) were performed as recommended by the supplier of the *Taq* polymerase (Perkin Elmer Cetus).

Amplifications of RNA by reverse transcription of the target RNA followed by standard DNA amplification

were performed using the Gene Amp RNA PCR Kit as recommended by the supplier (Perkin Elmer Cetus).

Examples

Example 1: Isolation of INSV particles and genetic material therein

INSV isolate NL-07, an isolate from *Impatiens*, is maintained on *Impatiens* by grafting. Virus is purified from systemically infected *Nicotiana rustica* leaves, after mechanical inoculation essentially as described by Tas et al. [(1977) J. Gen. Virol. 36:81-91]. All material used in the isolation procedure should be maintained at a temperature of 4 °C. Twelve days after inoculation 100 grams of infected leaves are harvested and ground for 5 - 10 seconds at a low speed setting in 5 volumes extraction buffer (0.1 M NaH₂PO₄, 0.01 M Na₂SO₃, pH 7) in a Waring blender. The suspension is filtered through cheesecloth and the filtrate is centrifuged for 10 minutes at 16,000 x g. The resulting pellet is resuspended in three volumes resuspension buffer (0.01 M NaH₂PO₄, 0.01 M Na₂SO₃, pH 7). The pellet is dissolved by stirring carefully at 4°C. After centrifuging for 10 minutes at 12,500 x g the pellet is discarded and the supernatant centrifuged again for 20 minutes at 50,000 x g. The pellet is resuspended in 0.2 volume of resuspension buffer (0.01 M NaH₂PO₄, 0.01 M Na₂SO₃, pH 7) and kept on ice for 30 minutes. Anti-serum raised in rabbits against material from non-infected *Nicotiana rustica* is added to the solution and carefully stirred for 1 hour. Non-viral complexes are pelleted after 10 minutes centrifuging at 16,000 x g. The cleared supernatant is loaded on a linear 5% - 40 % sucrose gradient in resuspension buffer (0.01 M NaH₂PO₄, 0.01 M Na₂SO₃, pH 7), and spun for 45 minutes at 95,000 x g. The opalescent band containing INSV particles is carefully collected with a syringe and diluted 4 times with resuspension buffer. Washed viruses are pelleted by centrifugation for 1.5 hours at 21,000 x g and resuspended in one volume of resuspension buffer. Generally, 100 grams of leaf material yields approximately 0.5 mg of INSV viruses. INSV RNA is recovered preferentially from purified virus preparations by SDS-phenol extractions followed by ethanol precipitation. From 1 mg INSV, 1-5 µg of RNA is extracted. The isolated RNA molecules are analysed for intactness by electrophoresis on an agarose gel. Three distinct RNA molecules are identified with apparent sizes of about 3000 nucleotides (S RNA), about 4900 nucleotides (M RNA) and about 8900 nucleotides (L RNA) respectively.

Example 2: Sequence determination of the 3'-termini of the INSV viral RNAs

In order to perform direct RNA sequencing, INSV RNA is extracted from purified nucleocapsids essentially according to Verkleij et al. (1983) supra. Twelve days after inoculation 100 grams of infected leaves are harvested and ground for 5 - 10 seconds at a low speed setting in four volumes of TAS-E buffer (0.01 M EDTA, 0.01 M Na₂SO₃, 0.1 % cysteine, 0.1 M TRIS pH 8.0) in a Waring blender. The suspension is filtered through cheesecloth and centrifuged for 10 minutes at 1,100 x g. Nucleocapsids are recovered from the supernatant after 30 minutes of centrifuging at 66,000 x g. The pellet is carefully resuspended in one volume of TAS-R buffer (1 % Nonidet NP-40, 0.01 M EDTA, 0.01 M Na₂SO₃, 0.1 % cysteine, 0.01 M glycine, 0.01 M TRIS, pH 7.9). The pellet is dissolved by stirring carefully for 30 minutes at 4 °C. The supernatant is cleared by centrifuging for 10 minutes at 16,000 x g. Crude nucleocapsids are collected from the cleared supernatant by sedimentation through a 30 % sucrose cushion for 1 hour at 105,000 x g. The nucleocapsid pellet is resuspended in 400 µl 0.01 M Na-citrate pH 6.5, layered on a 20 - 40 % sucrose (in 0.01 M Na-citrate pH 6.5) and spun for 2 hours at 280,000 x g. The three different opalescent bands, respectively L, M and S nucleocapsid, are collected separately. INSV RNA is recovered preferentially from purified nucleocapsid preparations by SDS-phenol extractions followed by ethanol precipitation. Generally, 100 µg of RNA are obtained from 100 grams of infected leaves. The 3'-ends of the separate INSV RNAs are labeled using RNA ligase and 5'-[³²P]pCp. The end-labeled RNA molecules are separated on a low gelling temperature agarose gel [Wieslander, (1979) Anal Biochem 98: 305-309]. The enzymatic approach described by Clerx-Van Haaster and Bishop [(1980) Virology 105:564-574] and Clerx-Van Haaster et al. [(1982) J Gen Virol 61:289-292] is used to determine the 30 terminal nucleotides of the 3'-and 5'-ends of both S and M RNA.

Synthetic oligonucleotides complementary to the 3'-termini are synthesized using a commercially available system (Applied Biosystems) and used for dideoxy-sequencing with reverse transcriptase.

Example 3: cDNA cloning of INSV genetic material

Oligonucleotides complementary to the 3'-end of the S RNA are used for priming first strand cDNA synthesis. With these primers, double stranded DNA to INSV RNA is synthesized according to Gubler and Hoffman [(1983) Gene 25:263-269].

Two different approaches are used to generate cDNA clones to the INSV viral RNAs. A first series of clones is obtained by random priming of the INSV RNA using fragmented single stranded calf thymus DNA, followed by first and second strand cDNA synthesis. cDNA is made blunt-ended using T4-DNA polymerase and ligated with T4 ligase into the Sma I site of pUC19.

A second series of INSV cDNA clones is obtained by priming first strand DNA synthesis with the oligonucleotides complementary to the 20 terminal nucleotides at the 3'-ends of the INSV RNAs. Blunt ended cDNA fragments are cloned into the Sma I site of pUC19.

cDNA clones from both series containing viral inserts are selected via colony hybridization, essentially according to the method of Grunstein and Hogness [(1975) Proc. Natl. Acad. Sci. USA 72:3961-3965] using [³²P]-labeled, randomly primed first strand cDNA as a probe. Sets of overlapping cDNA clones are selected by Southern analysis followed by plasmid walking, in order to construct a restriction map, based on cDNA derived sequences of the S RNA (Figure 2).

Example 4: Sequence determination of the INSV S RNA

In order to determine the sequence of the S RNA 5 selected cDNA clones are subcloned into pBluescript, resulting in the plasmids pINSV-S2, pINSV-S15, pINSV-S61, pINSV-S60 and pINSV-S39, (Figure 2). The clones are sequenced in both directions using the protocol of Zhang et al. [(1988) Nucl. Acids. Res. 16:1220]. The nucleotide sequence of the 3'-end of the S RNA is determined by primer extension of the synthetic oligonucleotide INSV-S60 (5' d(AGAGCAATTGTGTCA) which is complementary to the 15 nucleotides of the 3'-terminus. Sequence data from the INSV S RNA (3017 nt) is summarized in the sequence listing (SEQ ID No.1 to SEQ ID No.12).

Computer simulated translation of the 6 different reading frames on the viral strand and viral complementary strand reveals the presence of two putative open reading frames (Figure 3). On the viral strand an open reading frame is found starting at position 87 and terminating at a UAA stopcodon at position 1436 encoding a protein of 449 amino acids with a predicted molecular mass of about 51.2 kd. This protein is a non-structural protein, tentatively designated NSs (Figure 3/ SEQ ID No.26). The other open reading frame is located on the viral complementary strand from position 2080 to 2868 (SEQ ID No. 11), encoding a 262 amino acid long polypeptide with a predicted molecular mass of about 28.7 kd. This open reading frame encodes the viral nucleocapsid protein N (Figure 3/ SEQ ID No 25). Thus Figure 3 shows the coding capacities of the viral and the viral complementary strand of INSV S RNA, indicating the NSs and N protein genes are expressed from subgenomic mRNAs (SEQ ID No.3, SEQ ID No.11 respectively). Thus, the situation occurs that a plant virus RNA has an ambisense gene arrangement. Other important features of this S RNA sequence is the existence of complementary terminal repeats capable of forming so-called "pan-handle" structures. These structures play an important role in replication and transcription of viral RNA. Another putative regulatory element is the secondary structure in the intergenic region of the S RNA, which most likely contains the transcription termination signals for both subgenomic mRNAs, encoding respectively the N and NSs-protein.

The nucleotide sequence of the INSV M and L RNA is elucidated employing similar strategies and methods as used to determine the nucleotide sequence of the S RNA.

Example 5: Construction of an expression vector pZU-B

The recombinant plasmid pZO347 is a derivative of pBluescript carrying a 496 bp BamHI-SmaI fragment containing a 426 bp 35S promoter fragment (HincII fragment) of CaMV strain Cabb-S, linked to a 67 bp fragment of the non-translated leader region, the so-called Ω -region, of the tobacco mosaic virus. This results in a chimeric promoter with a complete transcriptional fusion between the promoter of CaMV to the untranslated leader of TMV. By using *in vitro* mutagenesis the original position of the TMV ATG startcodon is mutated to a SmaI site.

The plasmid pZO008 carries the nopaline synthase (NOS) terminator as a 260 bp PstI-HindIII fragment. This PstI-HindIII fragment is excised from pZO008 and ligated using T4 ligase into PstI-HindIII-linearized pZO347. The resulting recombinant plasmid pZU-B is another plant expression vector. The sequence of this 35S- Ω promoter as used in the plant expression vector pZU-B is shown as SEQ ID No.23. The resulting recombinant plasmid pZU-B contains the 35S HincII-TMV Ω fusion (35S- Ω), unique SmaI and PstI sites and the NOS terminator (Figure 4). This expression vector is preferentially used in constructing translational fusions of the gene for expression downstream of the chimaeric promoter 35S- Ω .

Example 6: Subcloning of the INSV N protein gene

The INSV N protein coding sequence is obtained by fusion of the cDNA clones pINSV-S60 and pINSV-S39 (Figure 5). The cDNA clone pINSV-S60 is subjected to *SpeI* digestion and the fragment containing the 3'-end of the INSV N protein gene is separated electrophoretically and purified from the gel using a DEAE membrane (NA-45, Schleicher and Schüll) and cloned in the largest *SpeI* fragment of pINSV-S39 linearized resulting in the recombinant plasmid pINSV-N. Primers are designed homologous to the translational start and stop codon. Primer INSV-066 d(GCAGATATCATGAACAAAGC) creates an *EcoRV* site just proximal to the start codon.

Primer INSV-070 d(GCAACCTGCAGCTCAAATCTCTT) creates a *PstI* site just distal to the stop codon. These primers are used in standard PCR experiments in which pINSV-N is used as the template. The resulting PCR fragment is isolated from the gel using a DEAE membrane (NA-45, Schleicher and Schüll) and cloned in the *SmaI* linearized pBluescript to generate plasmid pINSV-N2. The added restriction sites, *EcoRV* and *PstI*, facilitate the construction of further plasmids. (Alternatively, one may choose to add the sites in different ways such as but not limited to site-directed mutagenesis or by ligation of other synthetic oligonucleotide linkers. Such methods are all known to a person skilled in the art.)

Example 7: Subcloning of the INSV non-structural protein genes (NSs gene) of INSV S RNA

The sequence of the gene corresponding to the non-structural protein NSs is isolated using RNA based PCR on isolated INSV S RNA. Two primers are designed which are homologous to regions spanning either the translational start codon or stop codon. The start codon primer contains an *EcoRV* site proximal to the ATG codon, the stop codon primer has a *PstI* site just distal thereto. Purified INSV S RNA is subjected to the Gene AMP RNA PCR. The resulting PCR fragment is isolated from the gel and cloned into *SmaI* linearized pBluescript yielding the recombinant plasmid pINSV-NSs (Figure 6).

Example 8: Subcloning of the INSV non-structural protein gene (NSm gene) of the INSV M RNA

The sequence of the gene corresponding to the non-structural protein NSm is isolated using RNA based PCR on isolated INSV M RNA. Two primers are designed which are homologous to regions spanning either the translational start codon or stop codon. The start codon primer contains an *EcoRV* site proximal to the ATG codon, the stop codon primer has a *PstI* site just distal thereto. Purified INSV S RNA is subjected to the Gene AMP RNA PCR. The resulting PCR fragment is isolated from the gel and cloned into *SmaI* linearized pBluescript yielding the recombinant plasmid pINSV-NSm (Figure 7).

Example 9: Subcloning of the INSV G1/G2 glycoprotein gene (G1/G2 gene) of the INSV M RNA

The sequence of the gene corresponding to the G1/G2 glycoprotein precursor is isolated using RNA based PCR on isolated INSV M RNA. Two primers are designed homologous to regions spanning either the translational start codon or stop codon. The start codon primer contains an *EcoRV* site proximal to the ATG codon, the stop codon primer has a *PstI* site just distal thereto. Purified INSV M RNA is subjected to the Gene AMP RNA PCR. The resulting PCR fragment is isolated from the gel and cloned into *SmaI* linearized pBluescript yielding the recombinant plasmid pINSV-G1/G2 (Figure 8).

Example 10: Construction of plant transformation vectors containing INSV sequences**Example 10A: N protein constructions in pZU-B**

In order to make a fusion in which the ATG start codon from the N protein gene is fused directly to the 3'-end of the TMV untranslated leader of the 35S- Ω promoter the start codon of the N gene has to be mutated using the PCR approach as hereinbefore described. The N protein gene is excised from the plasmid pINSV-N2 via an *EcoRV*-*PstI* digestion. The fragment is isolated and inserted into the *SmaI*-*PstI* linearised pZU-B, resulting in recombinant plasmid pINSV-NB. The chimeric cassette containing the 35S- Ω promoter, the N gene and the NOS terminator is excised from the plasmid pINSV-NB via a *Bam*HI/*Xba*I digestion. The isolated chimeric gene cassette is then inserted into the *Bam*HI/*Xba*I linearized pBIN19 to create the binary transformation vector pINSV-NBB. The resulting plasmid pINSV-NBB (Figure 9) is used in plant transformation experiments using methods well known to a person skilled in the art.

Exempl 10B: NSs protein gene constructions in pZU-B

In order to create a fusion in which the ATG start codon from the NSs protein is fused directly to the 3'-end of the TMV leader of the 35S- Ω promoter the start codon of the NSs gene is mutated, using the PCR approach. The plasmid pINSV-Ns is digested with EcoRV and PstI and the NSs containing fragment is isolated from the gel and inserted into SmaI/PstI linearized pZU-B resulting in the recombinant plasmid pINSV-NSsB. The chimaeric cassette containing the 35S- Ω promoter, the mutated NSs protein gene and the NOS terminator is excised from the plasmid pINSV-NSsB via a BamHI/XbaI digestion. The isolated chimeric gene cassette is then inserted into the BamHI/XbaI linearized pBIN19 to create the binary transformation vector pINSV-NSsBB. The resulting plasmid pINSV-NSsBB (Figure 10) is used in plant transformation experiments using methods well known to a person skilled in the art.

Example 10C: G1/G2 glycoprotein gene constructions in pZU-B

In order to create a fusion in which the ATG start codon from the G1/G2 glycoprotein precursor is fused directly to the 3'-end of the TMV leader of the 35S- Ω promoter the start codon of the G1/G2 gene is mutated, using the PCR approach. The plasmid pINSV-G1/G2 is digested with EcoRV and PstI and the G1/G2 containing fragment is isolated from the gel and inserted into SmaI/PstI linearized pZU-B resulting in the recombinant plasmid pINSV-G1/G2B. The chimeric cassette containing the 35S- Ω promoter, the mutated G1/G2 glycoprotein gene and the NOS terminator is excised from the plasmid pINSV-G1/G2B via a BamHI/XbaI digestion. The isolated chimeric gene cassette is then inserted into the BamHI/XbaI linearized pBIN19 to create the binary transformation vector pINSV-G1/G2BB. The resulting plasmid pINSV-G1/G2BB (Figure 11) is used in plant transformation experiments using methods well known to a person skilled in the art.

Example 10D: NSm protein gene constructions in pZU-B

In order to create a fusion in which the ATG start codon from the NSm protein is fused directly to the 3'-end of the TMV leader of the 35S- Ω promoter the start codon of the NSm gene is mutated, using the PCR approach. The plasmid pINSV-NSm is digested with EcoRV and PstI and the NSm-containing fragment is isolated from the gel and inserted into SmaI/PstI linearized pZU-B resulting in the recombinant plasmid pINSV-NSmB. The chimeric cassette containing the 35S- Ω promoter, the mutated NSm protein gene and the NOS terminator is excised from the plasmid pINSV-NSmB via a BamHI/XbaI digestion. The isolated chimeric gene cassette is then inserted into the BamHI/XbaI linearized pBIN19 to create the binary transformation vector pINSV-NSmBB. The resulting plasmid pINSV-NSmBB (Figure 12) is used in plant transformation experiments using methods well known to a person skilled in the art.

Example 10E: 5'- and 3'-termini "pan-handle" constructions in pZU-B

A DNA analysis programme is used to locate the "pan-handle" element of the loop in the viral INSV S RNA. The strongest "pan-handle" structure that is detected includes about the first 24-25 nucleotides at the 5'-end (1 to 24 or 25) of the viral S RNA and about the last 36 nucleotides at the 3'-end of the viral S RNA (SEQ ID Nos 5 and 6 respectively). The length of the pan-handle element of the loop is about 36 nucleotides long.

These regions are synthesized on a commercial DNA synthesizer and appropriate linker sequences are added. Construction of the "pan-handle" vectors of S and M RNA results in respectively: pINSV-termS and pINSV-termM. Using appropriate restriction enzyme combination these fragments are inserted between the 35S- Ω promoter and the NOS terminator of pZU-B yielding the chimeric cassettes: pINSV-termSA, pINSV-termMA, pINSV-termSB and pINSV-termMB. These cassettes are then transferred into the binary transformation vector pBIN19 using appropriate enzyme combinations yielding the following plasmids: pINSV-termSAB, pINSV-termMAB, pINSV-termSBB and pINSV-termMBB. Alternatively, it is possible to design "pan-handle" constructs including the 3'- and 5'-end termini that are larger than indicated above, or separated by any other DNA sequence in order to enhance the stability of the transcripts produced from these recombinant genes in plants.

All "pan-handle" constructs resemble shortened tospovirus RNA molecules, specifically INSV RNA molecules and therefore can be regarded as defective interfering RNAs.

Exempl 10F: Construction containing INSV S RNA secondary structure region in pZU-B.

A DNA analysis programme is used to locate a secondary structure in the viral INSV S RNA. The strongest

secondary structure detectable starts at nucleotide 1440 and ends at nucleotide 2041 of SEQ ID No.1, (SEQ ID No 8).

The DNA fragment carrying the secondary structure region is isolated from pINSV-S61 using a PCR approach similar to that described earlier. The two primers used contain the sequences 1440-1460 and 2021-2041 of SEQ ID No.1. The PCR fragment is excised from an agarose gel and subsequently treated with T4 polymerase to create blunt ends and is subsequently cloned into the SmaI site of the expression vector pZU-B, resulting in the recombinant plasmid pINSV-HpSB. The plasmid pINSV-HpSB is digested with HindIII and the fragment containing the chimeric gene is excised from an agarose gel and ligated into XbaI linearized pBIN19, resulting in the transformation vector pINSV-HpSBB.

(It is clear to a person skilled in the art that other fragments can be isolated from the cDNA clones of the INSV S RNA containing the hairpin region as described above without interference to function. Also, a fragment containing the hairpin region may be synthesized using a DNA-synthesizer.)

15 **Example 11: Transformation of binary vectors to tobacco plant material**

Methods to transfer binary vectors to plant material are well established and known to a person skilled in the art. Variations in procedures exist due to for instance differences in used *Agrobacterium* strains, different sources of explant material, differences in regeneration systems depending on as well the cultivar as the plant species used.

The binary plant transformation vectors as described above are used in plant transformation experiments according to the following procedures. The constructed binary vector is transferred by tri-parental mating to an acceptor *Agrobacterium tumefaciens* strain, followed by southern analysis of the ex-conjugants for verification of proper transfer of the construct to the acceptor strain, inoculation and cocultivation of axenic explant material with the *Agrobacterium tumefaciens* strain of choice, selective killing of the *Agrobacterium tumefaciens* strain used with appropriate antibiotics, selection of transformed cells by growing on selective media containing kanamycine, transfer of tissue to shoot-inducing media, transfer of selected shoots to root inducing media, transfer of plantlets to soil, assaying for intactness of the construct by southern analyses of isolated total DNA from the transgenic plant, assaying for proper function of the inserted chimeric gene by northern analysis and/or enzyme assays and western blot analysis of proteins.

Example 12: Expression of INSV S RNA sequences in tobacco plant cells

RNA is extracted from leaves of regenerated plants using the following protocol. Grind 200 mg leaf material to a fine powder in liquid nitrogen. Add 800 µl RNA extraction buffer (100 mM Tris-HCl (pH 8,0), 500 mM NaCl, 2 mM EDTA, 200 mM β-Mercapto-ethanol, 0,4% SDS) and extract the homogenate with phenol, collect the nucleic acids by alcohol precipitation. Resuspend the nucleic acids in 0,5 ml 10 mM Tris-HCl (pH 8,0), 1 mM EDTA, add LiCl to a final concentration of 2 M, leave on ice for maximal 4 hours and collect the RNA by centrifugation. Resuspend in 400 µl 10 mM Tris-HCl (pH 8,0), 1 mM EDTA and precipitate with alcohol, finally resuspend in 50 µl 10 mM Tris-HCl (pH 8,0), 1 mM EDTA. RNAs are separated on glyoxal/agarose gels and blotted to Genescreen as described by van Grinsven et al. [(1986) Theor Appl Gen 73:94-101]. INSV S RNA sequences are detected using DNA or RNA probes labeled with [³²P], [³⁵S] or by using non-radioactive labeling techniques. Based on northern analysis, it is determined to what extent the regenerated plants express chimaeric INSV S RNA sequences.

Plants transformed with chimaeric constructs containing an INSV N protein-encoding sequence are also subjected to western blot analysis. Proteins are extracted from leaves of transformed plants by grinding in sample buffer according to the method of Laemmli [(1970) Nature 244: 29-30]. A 50 µg portion of protein is subjected to electrophoresis in a 12,5 % SDS-polyacrylamide gel essentially as described by Laemmli (1970) supra. Separated proteins are transferred to nitrocellulose electrophoretically as described by Towbin et al. [(1979) Proc. Natl. Acad. Sci. USA 76:4350-4354]. Transferred proteins are reacted with antiserum raised against purified INSV structural or non-structural proteins (Towbin et al.(1979) supra. Based on the results of the western analysis, it is determined that transformed plants do contain INSV N proteins encoded by the inserted chimaeric sequences.

55 **Example 13: Resistance of plants against INSV infection**

Transformed plants are grown in the greenhouse under standard quarantine conditions in order to prevent any infections by pathogens. The transformants are self-pollinated and the seeds harvested. Progeny plants are analyzed for segregation of the inserted gene and subsequently infected with INSV by mechanical inocu-

lation. Tissue from plants systemically infected with INSV is ground in 5 volumes of ice-cold inoculation buffer (10 mM phosphate buffer supplemented with 1% Na₂SO₃) and rubbed in the presence of carborundum powder on the first two fully extended leaflets of approximately 5 weeks old seedlings. Inoculated plants are monitored for symptom development during 3 weeks after inoculation.

Plants containing INSV Related DNA Sequences show reduced susceptibility to INSV infection as exemplified by a delay in symptom development, whereas untransformed control plants show severe systemic INSV symptoms within 7 days after inoculation.

10 **Example 14: Use of synthetic oligonucleotides for diagnostic purposes**

RNA is extracted from leaves of suspected plants using the following protocol: grind 1 gram of leaf material, preferentially showing disease symptoms, in 3 ml 100 mM Tris-HCl, 50 mM EDTA, 1.5 M NaCl and 2% CTAB (pH 8.0). After grinding, 1 ml of the homogenate is subjected to chloroform extraction and incubated at 65 °C for 10 minutes. The inorganic phase is then collected and extracted with phenol/chloroform (1:1), followed by a last extraction with chloroform. The ribonucleic acids are isolated from the inorganic phase, containing the total nucleic acids, by adding LiCl to a final concentration of 2 M. The preparation is incubated at 4°C for 1 hour, after which the ribonucleic acids are collected by centrifugation. The ribonucleic acid pellet is resuspended in 25 µl 10 mM Tris-HCl, 1 mM EDTA (pH 8.0). The ribonucleic acids are recovered by standard alcohol precipitation. The ribonucleic acid pellet is resuspended in 25 µl 10 mM Tris-HCl, 1 mM EDTA (pH 8.0).

1 µl of the purified ribonucleic acids is spotted on a nylon blotting membrane (e.g. Hybond-N, Amersham UK). The presence of INSV in the plant is detected by standard hybridization, using any part or parts of the sequence isolated from virions or preferentially by designing synthetic oligomers on the basis of disclosed sequence information as a probe. (Alternatively, in vitro transcripts of regions of the INSV genome are used to detect INSV Related RNA Sequences in diseased plants.) A diseased plant is diagnosed by the occurrence of hybridization at the dot containing RNA material from the diseased plant.

5

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

10

(A) NAME: SANDOZ LTD
(B) STREET: LICHTSTRASSE 35
(C) CITY: BASLE
(D) STATE: BS
(E) COUNTRY: SWITZERLAND
(F) POSTAL CODE (ZIP): CH-4002
(G) TELEPHONE: 061-3241111
(H) TELEFAX: 061-3227532
(I) TELEX: 965 050 55

15

(A) NAME: SANDOZ-PATENT-GMBH
(B) STREET: HUMBOLDSTRASSE 3
(C) CITY: LORRACH
(E) CCOUNTRY: GERMANY
(F) POSTAL CODE (ZIP): D-7850
(G) TELEPHONE: 076 21/4 80 14

20

(A) NAME: SANDOZ-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT
M.B.H.
(B) STREET: BRUNNER STRASSE 59
(C) CITY: VIENNA
(E) COUNTRY: AUSTRIA
(F) POSTAL CODE (ZIP): A-1230
(G) TELEPHONE: 86 93 61

25

(ii) TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO ORGANIC
COMPOUNDS

30

(iii) NUMBER OF SEQUENCES: 27

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

35

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: EP 93810190.4

40

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3017 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

45

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

50

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

55

5	AGAGCAATNN NNNNNNNNNN NNNNGAACAA CCAAGCTACA ACAAATCTTA CAATATTGTC	60
	AATTACATTA CTACTTCCAT TTTAACATGT CTAGTGCAAT GTATGAAACA ATTATCAAAT	120
10	CGAAGTCCTC AATCTGGGGA ACAACATCTT CGGCTAAAGC AGTAGTAGAT AGTTATTGGA	180
	TTCATGATCA ATCTTCCGGA AAGAAGTTGG TCGAAGCTCA ACTCTATTCT GACTCCAGGA	240
	GCAAGACCAG TTTCTGTTAC ACTGGTAAAG TTGGCTTTCT CCCAACAGAA GAAAAAGAAA	300
15	TTATAGTGAG ATGTTTTGIG CCTATTTTTG ATGACATTGA TCTGAATTC TCCTTTTCAG	360
	GGAATGTTGT CGAAATTCTG GTCAGATCTA ACACAACAAA CACAAACGGT GTTAAACATC	420
	AAGGTCATCT CAAAGTGTTA TCCTCTCAGT TGCTCAGAAT GCTTGAAGAG CAAATAGCAG	480
	TGCCTGAAAT TACTTCAAGA TTCGGTCTGA AAGAATCTGA CATCTTCCCT CCAAATAATT	540
20	TCATTGAAGC TGCAATAAA GGATCATTGT CTGTGTCAA AGAAGTCCTT TTTGATGTCA	600
	AGTATTCAAA CAACCAATCC ATGGGCAAAG TCAGTGTCTT TCTCCTACC AGAAGTGTTT	660
	ATGAATGGCT GTACACACTT AAGCCTGTTT TTAACCAATC CCAGACCAAC AACAGGACAG	720
25	TAAACACTTT GGCTGTAAAA TCACTGGCAA TGTCTGCAAC TTCTGATTTA ATGTCAGATA	780
	CTCATTGCTT TGTCAGGCTC AATAATAACA AGCCTTTTAA AATCAGCCTT TGGATGCGCA	840
	TCCCTAAAAA AATGAAATCA AACACATACA GCCGGTCTT CACCTGTCT GATGAATCTT	900
	CTCCTAAAGA GTATTATATA AGCATTCAAT GTCTTCCGAA TCACAACAAT GTTGAAACAG	960
30	TCATTGAATA TAACTTTGAT CAGTCAAACC TCTTCTTGAA TCAACTCCTT CTAGCAGTGA	1020
	TTCATAAAAT TGAGATGAAT TTTTCTGATC TAAAGAACC TTACAATGTT ATCCATGATA	1080
	TGTCGTATCC TCAAAGAATT GTTCATTAC TCTTGAAAT CCACACAGAA CTGCTCAAA	1140
35	CTGTCTGTGA CAGTGTTTCAG CAAGACATGA TTGTCTTCAC TATAAATGAG CCAGATCTAA	1200
	AGCCAAAAAA GTTTGAGCTA GGGAAAAAGA CTTTAAATTA TTCAGAAGAT GGTATGGGA	1260
	GAAATATTT CCTTCTCAG ACCTTGAAAA GTCTTCCGAG AAACTCACAA ACAATGTCTT	1320
40	ATTTGGATAG CATCCAGATG CCCGATTGGA AATTTGACTA TGCTGCAGGT GAAATAAAAA	1380
	TTTCTCCTAG ATCAGAGGAT GTTTTGAAAG CTATTTCTAA ATTAGATTTA AATTAACCTT	1440
	GGTTAAACTT GTCCCTAAGT AAAGTTTGTT TACATGCATT TAGATCAGAT TAAACAAATC	1500
	TAATAACAGA TAAACCAAAA ACAATCATAT GAAATAAATA AATAAACATA AAATATATAA	1560
45	AAAATACAAA AAAAATCATA AAATAAATAA AAACCAAAAA AGGATGGCCT TCGGGCACAA	1620
	TTTGGTTGCT TTAATAATGC TTTAAATGA ATGTATTAGT AAATTATAAA CTTTAAATCC	1680
	AATCTACTCA CAAATTGGCC AAAAATTTGT ATTTGTTTTT GTTTTGTGTT TTTGTTTTTT	1740
50	GTTTTTGTGTT TGTTTTATTT GTTTTTTATT TTGTTTTTTG TTTTTTGTGTT TTTATTTTAT	1800
	TTATATATAT ATATATATAT ATTTTGTAGT GGTTTTTATT GTTTTTATTA TTTTTGTAG	1860
	CTTTTTTACT TGTTTATTTT ACACGCAAAAC ACACTTTCAA GTTTATATAT TAAAACACAC	1920

5
 ATTAAACTTA TTTCAATAA TTTATAAAAG CACACTTAAT ACACTCAAAC AATAATTAAT 1980
 TATTTTATTT TTTATTTTAT TTTTATTTT TATTATTTT ATTTTATTT ATTTAAATGC 2040
 ATTTAACACA ACACAAAGCA AACCAAGCTC AAATCTCTT TAAATAGAAT CATTTTCCC 2100
 10 AAAATCAATA GTAGCATTAA ACATGCTGTA AATGGATGTA AGCCCTTCTT TGTAGTGGTC 2160
 CATTGCAGCA AGTCCTTCTAG CTTTCGGACT ACAAGCCTTT AGTATATCTG CATATTGTTT 2220
 AGCCTTGCCA ATTTCAACAG AGTTCATGCT ATATCCTTTG CTTTTTAGAA CTGTGCACAC 2280
 15 TTTCCCAACT GCCTCTTTAG TGCTAAACTT AGACATGTCA ATTCCAAGCT CAACATGTTT 2340
 AGCATCTTGA TAAATAGCCG GAACTAGTGC AGCTATTTC AATTCAGTA CAGATGCTAT 2400
 CAGAGGAAGA CTTCTCCTA AGAGAACACC CAAGACACAG GATTTCAAAT CTGTGGTTGC 2460
 20 AAGACCATAT GAGGCAATCA GAGGGTGACT TGGAAGGCTA TTTATAGCTT CAGTCAGAGC 2520
 AGATCCATTG TCCTTTATCA TTCCAACAAG ATGAACTCTC ACCATTGCAT CAAGTCTTCG 2580
 GAAAGTCATA TCATTGACCC CAACTCTTTC TGAATTGTTT CTAGTTTCT TAATTGTGAC 2640
 TGATCCAAAA GTGAAGTCAG CACTCTTAAT GACTCTCATT ATAGATTGCC TATTCTTGAG 2700
 25 GAAGGATAGG CAGGATGCAG TAGTCATGTT CTGAATCTT TCACGGTGT TGGTAAAGAA 2760
 GTCAGTGAAT TTGAAAGACC CTTCATTTTG AGTTTCCTCA AATTCTAAGG AATCAGATTG 2820
 AGTCAAAAGC TTGACTATGT TCTCCTTGGT AATCTTTGCT TTGTTTCTCT TGATCTGCTG 2880
 30 ACTTACTTAA CTTTAAAGCT TAAAGTGTC AAATTACTAA ATAGTACTTG CGGTAAAGT 2940
 AGTATTTGGT AAAATTTGTA ATTTTTCAGT TTCTAGCTT GGATTATATG ATGTTATATT 3000
 CGTGACACAA TTGCTCT 3017

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2993 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAACAACCAA GCTACAACAA ATCTTACAAT ATTGTCAATT ACATTACTAC TTCCATTTA 60
 50 ACATGTCTAG TGCAATGTAT GAAACAATTA TCAAATCGAA GTCCTCAATC TGGGGAACAA 120
 CATCTTCGGG TAAAGCAGTA GTAGATAGTT ATTGGATTCA TGATCAATCT TCCGGAAAGA 180

5	AGTTGGTCCG AGCTCAACTC TATTCTGACT CCAGGAGCAA GACCAGTTTC TGTTACACTG	240
	GTAAGTTGG CTTTCTCCCA ACAGAAGAAA AAGAAATTAT AGTGAGATGT TTTGTGCTTA	300
10	TTTTTGATGA CATTGATCTG AATTTCTCCT TTTCAGGGAA TGTTGTGCGA ATTCTGGTCA	360
	GATCTAACAC AACAAACACA AACGGTGTTA AACATCAAGG TCATCTCAAA GTGTTATCCT	420
	CTCAGTIGCT CAGAATGCTT GAAGAGCAAA TAGCAGTGCC TGAAATTACT TCAAGATTCTG	480
	GTCTGAAAGA ATCTGACATC TTCCCTCCAA ATAATTTTCAT TGAAGCTGCA AATAAAGGAT	540
15	CATTGTCTTG TGTCAAAGAA GTCCTTTTGT ATGTCAAGTA TTCAAACAAC CAATCCATGG	600
	GCAAAGTCAG TGTTCTTTCT CCTACCAGAA GTGTTTCATGA ATGGCTGTAC ACACTTAAGC	660
	CTGTTTTTAA CCAATCCCAG ACCAACAACA GGACAGTAAA CACTTTGGCT GTAAAATCAC	720
20	TGGCAATGTC TGCAACTTCT GATTTAATGT CAGATACTCA TTCGTTTGTC AGGCTCAATA	780
	ATAACAAGCC TTTTAAATC AGCCTTTGGA TGCGCATCCC TAAAATAATG AAATCAAACA	840
	CATACAGCCG GTTCTTCACC CTGTCTGATG AATCTTCTCC TAAAGAGTAT TATATAAGCA	900
25	TTCAATGTCT TCCGAATCAC AACAAATGTTG AAACAGTCAT TGAATATAAC TTTGATCAGT	960
	CAACCTCTT CTGGAATCAA CTCCTTCTAG CAGTGATTCA TAAAATTGAG ATGAATTTT	1020
	CTGATCTAAA AGAACCTTAC AATGTTATCC ATGATATGTC GTATCCTCAA AGAATTGTTC	1080
	ATTCACCTCT TGAAATCCAC ACAGAACTTG CTCAAAGTGT CTGTGACAGT GTTCAGCAAG	1140
30	ACATGATTGT CTTCACTATA AATGAGCCAG ATCTAAAGCC AAAAAAGTTT GAGCTAGGGA	1200
	AAAAGACTTT AAATTATTCA GAAGATGGTT ATGGGAGAAA ATATTTCTTT TCTCAGACCT	1260
	TGAAAGTCT TCCGAGAAAC TCACAAACAA TGTCTTATTT GGATAGCATC CAGATGCCCG	1320
35	ATTGGAAATT TGAATATGCT GCAGGTGAAA TAAAAATTTT TCCTAGATCA GAGGATGTTT	1380
	TGAAAGCTAT TTCTAAATTA GATTTAAATT AACCTTGGTT AAAGTTGTCC CTAAGTAAAG	1440
	TTTGTTTACA TGCATTTAGA TCAGATTAAA CAAATCTAAT AACAGATAAA CCAAAAACAA	1500
40	TCATATGAAA TAAATAAATA AACATAAAAT ATATAAAAAA TACAAAAAAA ATCATAAAAT	1560
	AAATAAAAC CAAAAAAGGA TGGCCTTCGG GCACAATTG GTTGCTTTAA TAATGCTTTA	1620
	AAATGAATGT ATTAGTAAAT TATAAACTTT AAATCCAATC TACTCACAAA TTGGCCAAAA	1680
	ATTTGTATTT GTTTTGTGTT TTGTTTTTTG TTTTGTGTT TTGTTTTGTT TTATTTGTTT	1740
45	TTTATTTTGT TTTTGTGTTT TTGTTTTTTA TTTTATTTAT ATATATATAT ATATATATTT	1800
	TGTAGTGGTT TTTATTGTTT TTATTATTTT TTGTAGCTTT TTTACTTGTT TATTTACAC	1860
	GCAACACAC TTTCAAGTTT ATATATTAAA ACACACATTA AACTTATTTT AAATAATTTA	1920
50	TAAAAGCACA CTTAATACAC TCAACAATA ATTAATTATT TTATTTTTTA TTTTATTTTT	1980
	TATTTTTATT ATTTTTATTT TTATTTATTT AAATGCATTT AACACAACAC AAAGCAAACC	2040
	AAGCTCAAT CTCTTTTAAA TAGAATCATT TTTCCAAAAA TCAATAGTAG CATTAAACAT	2100

5

GCTGTAAATG GATGTAAGCC CTTCTTTGTA GTGGTCCATT GCAGCAAGTC CTTTAGCTTT 2160
 CGGACTACAA GCCTTTAGTA TATCTGCATA TTGTTTAGCC TTGCCAATTT CAACAGAGTT 2220
 10 CATGCTATAT CCTTTGCTTT TTAGAACTGT GCACACTTTC CCAACTGCCT CTTTAGTGCT 2280
 AAACCTTAGAC ATGTCAATTC CAAGCTCAAC ATGTTTAGCA TCTTGATAAA TAGCCGGAAC 2340
 TAGTGCAGCT ATTTCAAAAT TCAGTACAGA TGCTATCAGA GGAAGACTTC CTCCTAAGAG 2400
 AACACCCAAG ACACAGGATT TCAAATCTGT GGTTCGAAGA CCATATGAGG CAATCAGAGG 2460
 15 GTGACTTGGA AGGCTATTTA TAGCTTCAGT CAGAGCAGAT CCATTGTCCT TTATCATTC 2520
 AACAAAGATGA ACTCTACCA TTGCATCAAG TCTTCGGAAA GTCATATCAT TGACCCCAAC 2580
 TCTTTCTGAA TTGTTTCTAG TTTTCTTAAT TGTGACTGAT CCAAAAGTGA AGTCAGCACT 2640
 20 CTTAAAGACT CTCATTATAG ATTGCCTATT CTTGAGGAAG GATAGGCAGG ATGCAGTAGT 2700
 CATGTCTGA ATCTTTTCAC GGTTGTTGGT AAAGAAGTCA GTGAAATTGA AAGACCCTTC 2760
 ATTTTGAGTT TCCTCAAAT CTAAGGAATC AGATTGAGTC AAAAGCTTGA CTATGTTCTC 2820
 CTEGGTAATC TTTGCTTTGT TCATCTTGAT CTGCTGACTT TACTAACTTT AAAGCTTAAA 2880
 25 GTGTTCAAAT TACTAAATAG TACTTGCGGT TAAAGTAGTA TTGGTAAAA TTTGTAATTT 2940
 TTCAGTTTCT AGCTTTGGAT TATATGATGT TATATTCGTG ACACAATTGC TCT 2993

(2) INFORMATION FOR SEQ ID NO:3:

30

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1350 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

35

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

40

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGTCTAGTG CAATGTATGA AACAAATTATC AAATCGAAGT CCTCAATCTG GGGAACAACA 60
 45 TCTTCGGGTA AAGCAGTAGT AGATAGTTAT TGGATTCTG ATCAATCTTC CGGAAAGAAG 120
 TTGGTCGAAG CTCAACTCTA TTCTGACTCC AGGAGCAAGA CCAGTTTCTG TTACTCTGGT 180
 AAAGTTGGCT TTCTCCCAAC AGAAGAAAAA GAAATTATAG TGAGATGTTT TGTGCCTATT 240
 TTTGATGACA TTGATCTGAA TTTCTCCTTT TCAGGGAATG TTGTCGAAAT TCTGGTCAGA 300
 50 TCTAACACAA CAAACACAAA CGGTGTTAAA CATCAAGGTC ATCTCAAAGT GTTATCCTCT 360
 CAGTTGCTCA GAATGCTTGA AGAGCAAATA GCAGTGCCTG AAATTACTTC AAGATTCGGT 420

55

5

CTGAAAGAAT CTGACATCTT CCCTCCAAAT AATTTCATTG AAGCTGCAAA TAAAGGATCA 480
 TTGTCTTGTG TCAAAGAAGT CCTTTTGTAT GICAAGTATT CAAACAACCA ATCCATGGGC 540
 10 AAAGTCAGTG TTCTTCTCC TACCAGAAGT GTTCATGAAT GGCTGTACAC ACTTAAGCCT 600
 GTTTTAAACC AATCCCAGAC CAACAACAGG ACAGTAAACA CTTTGGCTGT AAAATCACTG 660
 GCAATGTCTG CAACTTCTGA TTTAATGTCA GATACTCATT CGTTTGTGAG GCTCAATAAT 720
 AACAAGCCTT TTAAAAATCAG CCTTTGGATG CGCATCCCTA AAATAATGAA ATCAAACACA 780
 15 TACAGCCGGT TCTTCACCCT GTCTGATGAA TCTTCTCCTA AAGAGTATTA TATAAGCATT 840
 CAATGTCTTC CGAATCACAA CAATGTTGAA ACAGTCATTG AATATAACTT TGATCAGTCA 900
 AACCTCTTCT TGAATCAACT CCTTCTAGCA GTGATTCATA AAATGAGAT GAATTTTCT 960
 20 GATCTAAAAG AACCTTACAA TGTATCCAT GATATGTCGT ATCCTCAAAG AATTGTTTAT 1020
 TCACTTCTTG AAATCCACAC AGAACTTGCT CAACTGTCT GTGACAGTGT TCAGCAAGAC 1080
 ATGATTGTCT TCACTATAAA TGAGCCAGAT CTAAAGCCAA AAAAGTTTGA GCTAGGGAAA 1140
 AAGACTTTAA ATTATTCAGA AGATGGTTAT GGGAGAAAAT ATTCCTTTC TCAGACCTTG 1200
 25 AAAAGTCTTC CGAGAACTC ACAAACAATG TCTTATTGG ATAGCATCCA GATGCCCGAT 1260
 TGGAAATTG ACTATGCTGC AGGTGAAATA AAAATTTCTC CTAGATCAGA GGATGTTTGT 1320
 AAAGCTATTT CTAAATTAGA TTAAATTAA 1350

30

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 789 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 35 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

40

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

45 TTAAATAGAA TCATTTTCC CAAATCAAT AGTAGCATT AACATGCTGT AAATGGATGT 60
 AAGCCCTTCT TTGTAGTGGT CCATTGCAGC AAGTCCTTTA GCTTTCGGAC TACAAGCCTT 120
 TAGTATATCT GCATATTGTT TAGCCTTGCC AATTTCACA GAGTTCATGC TATATCCTTT 180
 50 GCTTTTLAGA ACTGTGCACA CTTCCCAAC TGCTCTTTA GTGCTAAACT TAGACATGTC 240
 AATTCCAAGC TCAACATGTT TAGCATCTTG ATAAATAGCC GGAAGTAGTG CAGCTATTTT 300
 AAAATTCAGT ACAGATGCTA TCAGAGGAAG ACTTCTCCT AAGAGAACAC CCAAGACACA 360

55

5

GGATTTCAAA TCTGTGGTTG CAAGACCATA TGAGGCAATC AGAGGGTGAC TTGGAAGGCT 420
 ATTTATAGCT TCAGTCAGAG CAGATCCATT GTCCTTTATC ATTCCAACAA GATGAAGTCT 480
 CACCATTCGA TCAAGTCTTC GGAAAGTCAT ATCATTGACC CCAACTCTTT CTGAATTGTT 540
 TCTAGITTTT TTAATTGTGA CTGATCCAAA AGTGAAGTCA GCACTCTTAA TGAATCTCAT 600
 TATAGATTGC CTATTCTTGA GGAAGGATAG GCAGGATGCA GTAGTCATGT TCTGAATCTT 660
 TTCACGGTTG TTGGTAAAGA AGTCAGTGAA ATTGAAAGAC CCTTCATTTT GAGTTTCCTC 720
 AAATTCTAAG GAATCAGATT GAGTCAAAAG CTTGACTATG TTCTCCTTGG TAATCTTTGC 780
 TTTGTTTCA 789

(2) INFORMATION FOR SEQ ID NO:5:

20

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

25

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO

30

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGAGCAATNN NNNNNNNNNN NNNNGAACAA CCCAAGC

37

35

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

40

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iii) ANTI-SENSE: NO

45

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GATTATATGA TGTTATATTC GTGACACAAT TGCTCT

36

50

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 643 base pairs

55

5

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

10

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

15

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

20

CCTTGCTTAA	ACTTGTCCT	AAGTAAAGTT	TGTTTACATG	CATTAGATC	AGATTAAACA	60
AATCTAATAA	CAGATAAACC	AAAAACAATC	ATATGAAATA	AATAAATAAA	CATAAAATAT	120
ATAAAAAATA	CAAAAAAAT	CATAAAATAA	ATAAAAACCA	AAAAAGGATG	GCCTTCGGGC	180
ACAATTTGGT	TGCTTTAATA	ATGCTTTTAA	ATGAATGTAT	TAGTAAATTA	TAAACTTTAA	240
ATCCAATCTA	CTCACAAATT	GGCCAAAAAT	TTGTATTTGT	TTTTGTTTTT	GTTTTTTGTT	300
TTTTGTTTTT	GTCTTGTCTT	ATTTGTTTTT	TATTTTGTTT	TTTGTTTTTT	GTTTTTTATT	360
TTATTTATAT	ATATATATAT	ATATATTTTG	TAGTGGTTTT	TATTGTTTTT	ATTATTTTTT	420
GTAGCTTTTT	TACTTGTTTA	TTTCACACGC	AAACACACTT	TCAAGTTTAT	ATATTAAAAC	480
ACACATTAAA	CTTATTTCAA	ATAATTTATA	AAAGCACACT	TAATACACTC	AAACAATAAT	540
TAATTATITT	ACTTTTTATT	TTATTTTTTA	TTTTTATTAT	TTTATTTTTT	ATTTATTTAA	600
ATGCATTTAA	CACAACACAA	AGCAAACCAA	GCTCAAATCT	CTT		643

35

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

40

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

50

TGCTTAAACT	TGTCCTAAG	TAAAGTTTGT	TTACATGCAT	TTAGATCAGA	TTAAACAAAT	60
CTAATAACAG	ATAAACCAAA	AACAATCATA	TGAATAAAAT	AAATAAACAT	AAAATATATA	120
AAAAATACAA	AAAAAATCAT	AAAATAAATA	AAAACCAAAA	AAGGATGGCC	TTCGGGCACA	180

55

5

10

15

ATTTGGTTGC TTTAATAATG CTTTAAAATG AATGTATTAG TAAATTATAA ACTTTAAATC 240
 CAATCTACTC ACAAATTGGC CAAAAATTG TATTTGTTTT TGTTTTGTGTT TTTTGTTTTT 300
 TGTTTTTGTT TTGTTTTATT TGTTTTTTAT TTTGTTTTTT GTTTTTTGTGTT TTTTATTETA 360
 TTTATATATA TATATATATA TATTCTGTAG TGGTTTTTAT TGTTTTTTATT ATTTTTTGTA 420
 GCTTTTTTAC TTGTTTATTT CACACGCAA CACACTTCA AGTTTATATA TTAAACACA 480
 CATTAAACTT ATTTCAAATA ATTTATAAAA GCACACTTAA TACACTCAA CAATAATTAA 540
 TTATTTTATT TTTTATTETA TTTTATTATT TTATTATTTT TATTTTATT TATTTAAATG 600
 CA 602

(2) INFORMATION FOR SEQ ID NO:9:

20

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3017 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

25

- (ii) MOLECULE TYPE: cDNA
 (iii) HYPOTHETICAL: NO
 (iii) ANTI-SENSE: YES

30

- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

35

40

45

50

AGAGCAATTG TGTCACGAAT ATAACATCAT ATAATCCAAA GCTAGAAACT GAAAAATTAC 60
 AAATTTTACC AAATACTACT TTAACCGCAA GTACTATTTA GTAATTTGAA CACTTTAAGC 120
 TTTAAAGTTA GTAAAGTCAG CAGATCAAGA TGAACAAAGC AAAGATTACC AAGGAGAACA 180
 TAGTCAAGCT TTTGACTCAA TCTGATTCTT TAGAATTTGA GGAAACTCAA AATGAAGGGT 240
 CTTTCAATTT CACTGACTTC TTTACCAACA ACCGTGAAAA GATTCAGAAC ATGACTACTG 300
 CATCCTGCCT ATCCTTCCTC AAGAATAGGC AATCTATAAT GAGAGTCATT AAGAGTGCTG 360
 ACTTCACTTT TGGATCAGTC ACAATTAAGA AAAGTAGAAA CAATTCAGAA AGAGTTGGGG 420
 TCAATGATAT GACTTTCCGA AGACTTGATG CAATGGTGAG AGTTCATCTT GTTGAATGA 480
 TAAAGGACAA TGGATCTGCT CTGACTGAAG CTATAAATAG CCTTCCAAGT CACCTCTGA 540
 TTGCCTCATA TGGTCTTGCA ACCACAGATT TGAAATCCTG TGTCTTGGGT GTTCTCTTAG 600
 GAGGAAGTCT TCCTCTGATA GCATCTGTAC TGAATTTTGA AATAGCTGCA CTAGTTCCGG 660
 CTATTTATCA AGATGCTAAA CATGTTGAGC TTGGAATTGA CATGTCTAAG TTTAGCACTA 720
 AAGAGGCAGT TGGGAAAGTG TGCACAGTTC TAAAAAGCAA AGGATATAGC ATGAACTCTG 780
 TTGAAATTGG CAAGGCTAAA CAATATGCAG ATATACTAAA GGCTTGTAAGT CCGAAAGCTA 840

55

5	AAGGACTTGC	TGCAATGGAC	CACTACAAAG	AAGGGCTTAC	ATCCATTAC	AGCATGTTTA	900
	ATGCTACTAT	FGATTTTGGG	AAAAATGATT	CTATTTAAAA	GAGATTTGAG	CTTGGTTTGC	960
10	TTTGTGTTGT	GTAAATGCA	TTTAAATAAA	TAAAAATAAA	AATAATAAAA	ATAAAAAATA	1020
	AAATAAAAAA	TAAATAATT	AATTATGTT	TGAGTGTATT	AAGTGTGCTT	TTATAAATTA	1080
	TTTGAAATAA	GTTTAATGIG	TGTTTAAATA	TATAAACTTG	AAAGTGTGCT	TGCGTGTGAA	1140
	ATAACAAGT	AAAAAAGCTA	CAAAAAATAA	TAAAAACAAT	AAAAACCACT	ACAAAATATA	1200
15	TATATATATA	TATATAAATA	AAATAAAAAA	CAAAAAACAA	AAAACAAAAT	AAAAAACAAA	1260
	TAAAAACAAA	CAAAAAACAA	AAACAAAAAA	CAAAAAACAA	AACAAATACA	AATTTTGGC	1320
	CAATTTGTGA	GTAGATTGGA	TTTAAAGTTT	ATAATTTACT	AATACATTCT	TTTAAAGCAT	1380
20	TATTAAAGCA	ACCAAATTGT	GCCCGAAGGC	CATCCTTTTT	TGGTTTTTAT	TTATTTTATG	1440
	ATTTTTTTTG	TATTTTTTAT	ATATTTTATG	TTTATTTATT	TATTTCATAT	GATTGTTTTT	1500
	GGTTTATCTG	TTATTAGATT	TGTTTAATCT	GATCTAAATG	CATGTAAACA	AACTTTACTT	1560
25	AGGGACAAGT	TTAACCAAGG	TTAATTTAAA	TCTAATTTAG	AAATAGCTTT	CAAAACATCC	1620
	TCTGATCTAG	GAGAAATTTT	TATTTACCT	GCAGCATAGT	CAAATTTCCA	ATCGGGCATC	1680
	TGGATGCTAT	CCAAATAAGA	CATTGTTTGT	GAGTTTCTCG	GAAGACTTTT	CAAGGTCTGA	1740
	GAAAGGAAAT	ATTTTCTCCC	ATAACCATCT	TCTGAATAAT	TTAAAGTCTT	TTTCCCTAGC	1800
30	TCAAACCTTT	TTGGCTTTAG	ATCTGGCTCA	TTTATAGTGA	AGACAATCAT	GTCTTGCTGA	1860
	ACACTGTCAC	AGACAGTTTG	AGCAAGTTCT	GTGTGGATT	CAAGAAGTGA	ATGAACAATT	1920
	CTTTGAGGAT	ACGACATATC	ATGGATAACA	TTGTAAGGTT	CTTTTAGATC	AGAAAAATTC	1980
35	ATCTCAATTT	TATGAATCAC	TGCTAGAAGG	AGTTGATTCA	AGAAGAGGTT	TGACTGATCA	2040
	AAGTTATATT	CAATGACTGT	TTCAACATTG	TTGTGATTCT	GAAGACATTG	AATGCTTATA	2100
	TAATACTCTT	TAGGAGAAGA	TTCATCAGAC	AGGGTGAAGA	ACCGGCTGTA	TGTGTTTGAT	2160
40	TTTATTATTT	TAGGGATGCG	CATCCAAAGG	CTGATTTTAA	AAGGCTTGTT	ATTATTGAGC	2220
	CTGACAAACG	AATGAGTATC	TGACATTAAA	TCAGAAGTTG	CAGACATTGC	CAGTGATTTT	2280
	ACAGCCAAAG	TGTTTACTGT	CCTGTTGTTG	GTCTGGGATT	GGTTAAAAAC	AGGCTTAAGT	2340
	GTGTACAGCC	ATTCATGAAC	ACTTCTGGTA	GGAGAAAGAA	CACTGACTTT	GCCCATGGAT	2400
45	TGGTTGTTTG	AATACTTGAC	ATCAAAAAGG	ACTTCTTTGA	CACAAGACAA	TGATCCTTTA	2460
	TTTGCAGCTT	CAATGAAATT	ATTTGGAGGG	AAGATGTCAG	ATTCTTTCAG	ACCGAATCTT	2520
	GAAGTAATTT	CAGGCACTGC	TATTTGCTCT	TCAAGCATT	TGAGCAACTG	AGAGGATAAC	2580
50	ACTTTGAGAT	GACCTTGATG	TTTAACACCG	TTTGTGTTTG	TTGTGTTAGA	TCTGACCAGA	2640
	ATTTGACAAA	CATTCCTGTA	AAAGGAGAAA	TTTCAAGTCAA	TGTCATCAAA	AATAGGCACA	2700
	AAACATCTCA	CTATAATTT	TTTTTCTTCT	GTTGGGAGAA	AGCCAACTTT	ACCAGTGTA	2760

5

CAGAACTGG TCTTGCTCCT GGAGTCAGAA TAGAGTTGAG CTTCGACCAA CTTCTTTCCG 2820
 GAAGATTGAT CATGAATCCA ATAATATCT ACTACTGCTT TACCCGAAGA TGTTGTTCCC 2880
 CAGATTGAGG ACTTCGATTT GATAATTGTT TCATACATTG CACTAGACAT GTTAAATGG 2940
 10 AAGTAGTAAT GTAATTGACA ATATTGTAAG ATTTGTTGTA GCTTGGTTGT TCNNNNNNNN 3000
 NNNNNNNNNA TTGCTCT 3017

(2) INFORMATION FOR SEQ ID NO:10:

15

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2993 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

20

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: YES

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGAGCAATTG TGTCACGAAT ATAACATCAT ATAATCCAAA GCTAGAACT GAAAAATTAC 60
 30 AAATTTTACC AAATACTACT TTAACCGCAA GTACTATTTA GTAATTTGAA CACTTTAAGC 120
 TTTAAAGTTA GTAAAGTCAG CAGATCAAGA TGAACAAAGC AAAGATTACC AAGGAGAACA 180
 TAGTCAAGCT TTTGACTCAA TCTGATTCCT TAGAATTTGA GGAACTCAA AATGAAGGGT 240
 35 CTTTCAATTT CACTGACTTC TTTACCAACA ACCGTGAAAA GATTCAGAAC ATGACTACTG 300
 CATCCTGCCT ATCCTTCCTC AAGAATAGGC AATCTATAAT GAGAGTCATT AAGAGTGCTG 360
 ACTTCACTTT TGGATCAGTC ACAATTAAGA AACTAGAAA CAATTCAGAA AGAGTTGGGG 420
 TCAATGATAT GACTTTCCGA AGACTTGATG CAATGGTGAG AGTTCATCTT GTTGGAAATGA 480
 40 TAAAGGACAA TGGATCTGCT CTGACTGAAG CTATAAATAG CCTTCCAAGT CACCCTCTGA 540
 TTGCCTCATA TGGTCTTGCA ACCACAGATT TGAAATCCTG TGTCTTGGGT GTTCTCTTAG 600
 GAGGAAGTCT TCCTCTGATA GCATCTGTAC TGAATTTTGA AATAGCTGCA CTAGTTCCGG 660
 45 CTATTTATCA AGATGCTAAA CATGTTGAGC TTGGAATTGA CATGTCTAAG TTTAGCACTA 720
 AAGAGGCAGT TGGGAAAGTG TGCACAGTTC TAAAAAGCAA AGGATATAGC ATGAACTCTG 780
 TTGAAATTGG CAAGGCTAAA CAATATGCAG ATATACTAAA GGCTTGTAAGT CCGAAAGCTA 840
 AAGGACTTGC TGCAATGGAC CACTACAAAG AAGGGCTTAC ATCCATTTAC AGCATGTTTA 900
 50 ATGCTACTAT TGATTTTGGG AAAAATGATT CTATTTAAAA GAGATTTGAG CTTGGTTTGC 960
 TTTGTGTTGT GTTAAATGCA TTAAATAAAA TAAAAATAAA AATAATAAAA ATAAAAAATA 1020

55

5	AAATPAAAAA TAAAATAATT AATTATTGTT TGAGTGTATT AAGTGTGCTT TTATAAATTA	1080
	TTTGAAATAA GTTTAATGIG TGTTTTAATA TATAAAGTTG AAAGTGTGTT TGCCTGTGAA	1140
10	ATAAACCAAGT AAAAAAGCTA CAAAAAATAA TAAAAACAAT AAAAACCACT ACAAATATA	1200
	TATATATATA TATATAAATA AAATAAAAAA CAAAAACAA AAAACAAAAT AAAAAACAA	1260
	TAAACAAAA CAAAAACAA AAACAAAAA CAAAAACAA AACAAATACA AATTTTGGC	1320
15	CAATTTGTGA GTAGATTGGA TTTAAAGTTT ATAATTTACT AATACATTCA TTTTAAAGCA	1380
	TTATTAAAGC AACCAAATG TGCCCGAAGG CCATCCTTTT TTGGTTTTTA TTTATTTTAT	1440
	GATTTTTTTT GTATTTTTTA TATATTTTAT GTTTATTTAT TTATTTTATA TGATTGTTTT	1500
	TGGTTTATCT GTTATTAGAT TTGTTTAAAT TGATCTAAAT GCATGTAAAC AAACCTTTACT	1560
20	TAGGGACAAG TTTAACCAAG GTTAATTTAA ATCTAATTTA GAAATAGCTT TCAAAACATC	1620
	CTCTGATCTA GGAGAAATTT TTATTTTACC TGACGATAG TCAAATTTCC AATCGGGCAT	1680
	CTGGATGCTA TCCAAATAAG ACATTGTTTG TGAGTTTCTC GGAAGACTTT TCAAGGTCTG	1740
25	AGAAAGGAAA TATTTTCTCC CATAACCATC TCTGAATAA TTTAAAGTCT TTTTCCCTAG	1800
	CTCAAATTTT TTTGGCTTTA GATCTGGCTC ATTTATAGTG AAGACAATCA TGTCTTGCTG	1860
	AACACTGTCA CAGACAGTTT GAGCAAGTTC TGTGTGGATT TCAAGAAGTG AATGAACAAT	1920
	TCTTTGAGGA TACGACATAT CATGGATAAC ATTGTAAGGT TCTTTTAGAT CAGAAAAATT	1980
30	CATCTCAATT TATGAATCA CTGCTAGAAG GAGTTGATTC AAGAAGAGGT TTGACTGATC	2040
	AAAGTTATAT TCAATGACTG TTTCAACATT GTTGTGATTC GGAAGACATT GAATGCTTAT	2100
	ATAATACTCT TTAGGAGAAG ATTCATCAGA CAGGGTGAAG AACCGGCTGT ATGTGTTTGA	2160
35	TTTCATTATT TTAGGGATGC GCATCCAAAG GCTGATTTTA AAAGGCTTGT TATTATTGAG	2220
	CCTGACAAAC GAATGAGTAT CTGACATTAA ATCAGAAGTT GCAGACATTG CCAGTGATT	2280
	TACAGCCAAA GTGTTTACTG TCCTGTTGTT GGTCTGGGAT TGGTTAAAAA CAGGCTTAAG	2340
40	TGTGTACAGC CATTATGAA CACTTCTGGT AGGAGAAAGA AACTGACTT TGCCCATGGA	2400
	TTGGTTGTTT GAATACTTGA CATCAAAAAG GACTTCTTTG ACACAAGACA ATGATCCTTT	2460
	ATTTGCAGCT TCAATGAAAT TATTTGGAGG GAAGATGTCA GATTCTTTCA GACCGAATCT	2520
	TGAAGTAATT TCAGGCACTG CTATTTGCTC TTCAAGCATT CTGAGCAACT GAGAGGATAA	2580
45	CACTTTGAGA TGACCTTGAT GTTTAACACC GTTTGTGTTT GTTGTGTTAG ATCTGACCAG	2640
	AATTTGACA ACATTCCCTG AAAAGGAGAA ATTCAGATCA ATGTCATCAA AAATAGGCAC	2700
	AAAACATCTC ACTATAATTT CTTTTCTTC TGTGGGAGA AAGCCAACTT TACCAGTGTA	2760
50	ACAGAACTG GTCTTGCTCC TGGAGTCAGA ATAGAGTTGA GCTTCGACCA ACTTCTTTCC	2820
	GGAAGATTGA TCATGAATCC AATACTATC TACTACTGCT TTACCCGAAG ATGTTGTTCC	2880
	CCAGATTGAG GACTTCGATT TGATAATTGT TTCATACATT GCACTAGACA TGTAAAAATG	2940

5

GAAGTAGTAA TGTAATTGAC AATATTGTAA GATTTGTTGT AGCTTGGTTG TTC

2993

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 789 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATGAACAAAG CAAAGATTAC CAAGGAGAAC ATAGTCAAGC TTTGACTCA ATCTGATTCC	60
TTAGAATTTG AGGAAACTCA AAATGAAGGG TCTTTCAATT TCACTGACTT CTTTACCAAC	120
AACCGTGAAA AGATTGAGAA CATGACTACT GCATCCTGCC TATCCTTCCT CAAGAATAGG	180
CAATCTATAA TGAGAGTCAT TAAGAGTGCT GACTTCACTT TTGGATCAGT CACAATTAAG	240
AAAAC TAGAA ACAATTCAGA AAGAGTTGGG GTCAATGATA TGACTTTCCG AAGACTTGAT	300
GCAATGGTGA GAGTTCATCT TGTTGGAATG ATAAAGGACA ATGGATCTGC TCTGACTGAA	360
GCTATAAATA GCCTTCCAAG TCACCCTCTG ATTGCCTCAT ATGGTCTTGC AACCACAGAT	420
TTGAAATCCT GTGTCTTGGG TGTTCTCTTA GGAGGAAGTC TTCCTCTGAT AGCATCTGTA	480
CTGAATTTTG AAATAGCTGC ACTAGTTCCG GCTATTTATC AAGATGCTAA ACATGTTGAG	540
CTTGGAATTG ACATGTCTAA GTTAGCACT AAAGAGGCAG TTGGGAAAGT GTGCACAGTT	600
CTAAAAAGCA AAGGATATAG CATGAACCTCT GTTGAAATTG GCAAGGCTAA ACAATATGCA	660
GATATACTAA AGGCTTGTAG TCCGAAAGCT AAAGGACTTG CTGCAATGGA CCACTACAAA	720
GAAGGGCTTA CATCCATTTA CAGCATGTTT AATGCTACTA TTGATTTTGG GAAAAATGAT	780
TCTATTTAA	789

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1350 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: YES

55

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTAATTTAAA TCTAATTTAG AAATAGCTTT CAAAACATCC TCTGATCTAG GAGAAATTTT	60
TATTTACCTT GCAGCATAGT CAAATTTCCA ATCGGGCATC TGGATGCTAT CCAAATAAGA	120
CATTGTTTGT GAGTTTCTCG GAAGACTTTT CAAGGTCTGA GAAAGGAAAT ATTTTCTCCC	180
ATAACCATCT TCTGAATAAT TTAAAGTCIT TTTCCCTAGC TCAAACCTTTT TTGGCTTTAG	240
ATCTGGGCTCA TTTATAGTGA AGACAATCAT GTCTTGCTGA ACACTGTCAC AGACAGTTTG	300
AGCAAGTTCT GTGTGGATTT CAAGAAGTGA ATGAACAATT CTTTGAGGAT ACGACATATC	360
ATGGATAACA TTGTAAGGTT CTTTLAGATC AGAAAAATTC ATCTCAATTT TATGAATCAC	420
TGCTAGAAGG AGTTGATTCA AGAAGAGGTT TGACTGATCA AAGTTATATT CAATGACTGT	480
TTCAACATTG TTGTGATTCT GAAGACATTG AATGCTTATA TAATACTCTT TAGGAGAAGA	540
TTCATCAGAC AGGGTGAAGA ACCGGCTGTA TGTGTTTGAT TTCATTATTT TAGGGATGCG	600
CATCCAAAGG CTGATTTTAA AAGGCTTGTT ATTATTGAGC CTGACAAACG AATGAGTATC	660
TGACATTAAA TCAGAAGTTG CAGACATTGC CAGTGATTTT ACAGCCAAAG TGTTTACTGT	720
CCTGTGTGTT GTCTGGGATT GGTAAAAAAC AGGCTTAAGT GTGTACAGCC ATTCATGAAC	780
ACTTCTGGTA GGAGAAAGAA CACTGACTTT GCCCATGGAT TGGTTGTTTG AATACTTGAC	840
ATCAAAAAGG ACTTCTTTGA CACAAGACAA TGATCCTTA TTTGCAGCTT CAATGAAATT	900
ATTTGAGGGG AAGATGTCAG ATTCTTTTCA ACCGAATCTT GAAGTAATT CAGGCACTGC	960
TATTTGCTCT TCAAGCATTG TGAGCAACTG AGAGGATAAC ACTTTGAGAT GACCTTGATG	1020
TTTAACACCG TTTGTGTTTG TTGTGTTAGA TCTGACCAGA ATTCGACAA CATTCCCTGA	1080
AAAGGAGAAA TTCAGATCAA TGTCATCAA AATAGGCACA AAACATCTCA CTATAATTTT	1140
TTTTTCTTCT GTTGGGAGAA AGCCAACTTT ACCAGTGTA CAGAACTGG TCTTGCTCCT	1200
GGAGTCAGAA TAGAGTTGAG CTTGACCAA CTTCTTTCCG GAAGATTGAT CATGAATCCA	1260
ATAACTATCT ACTACTGCTT TACCCGAAGA TGTGTGTTCC CAGATTGAGG ACTTCGATTT	1320
GATAATTGTT TCATACATTG CACTAGACAT	1350

45

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 642 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

50

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

55

5

(iii) ANTI-SENSE: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAGAGATTG AGCTTGTTT GCTTTGTGTT GTGTTAATG CATTAAATA AATAAAATA 60
 AAAATAATAA AAATAAAAAA TAAATAAAA AATAAAATAA TTAATTATTG TTTGAGTGTA 120
 15 TTAAGTGTGC TTTTATAAAT TATTTGAAAT AAGTTAATG TGTGTTTAA TATATAAACT 180
 TGAAAGTGTG TTTGCGTGTG AAATAAACAA GTAAAAAGC TACAAAAAT AATAAAAACA 240
 ATAAAAACCA CTACAAATA TATATATATA TATATATAA TAAATAAAA AACAAAAAAC 300
 20 AAAAAACAAA ATAAAAACA AATAAACAA AACAAAAACA AAAAAACAAA AACAAAAACA 360
 AAAACAAATA CAAATTTTGT GCCAATTGT GAGTAGATTG GATTTAAAGT TTATAATTTA 420
 CTAATACATT CTTTAAAGC ATTATTAAAG CAACCAAATT GTGCCGAAG GCCATCCTTT 480
 TTTGGTTTTT ATTTATTTTA TGATTTTTTT TGTATTTTTT ATATATTTTA TGTTTATTTA 540
 25 TTTATTTTAT ATGATTGTTT TTGTTTATC TGTTATTAGA TTTGTTTAA CTGATCTAAA 600
 TGCATGTAAA CAACTTTTAC TTAGGGACAA GTTTAACCA GG 642

(2) INFORMATION FOR SEQ ID NO:14:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4970 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

35

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

40

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGAGCAATCA GTGCATCAA ATTATATCTA GCCGAATTCA ATCATTATCT TCTCAATATT 60
 45 TTAATTCTTA ATCTACCGTC CAGAGATGAA TAGTTTTTTC AAATCACTCA GATCATCTAG 120
 CAGCAGGGAG CTAGATCACC CTAGGGTTAC AACTACCCTC TCTAAACAAG GAGCAGACAT 180
 TGTTGTACAC AATCCTCTG CTAATCACAA CAACAAGGAA GTTCTCCAAA GAGCCATGGA 240
 TAGCTCTAAA GGGAAGATTT TGATGAACAA TACAGGCACC TCATCACTAG GCACATATGA 300
 50 GTCTGACCAG ATATCTGAAT CAGAGTCTTA TGATCTTCT GCTAGAATGA TTGTTGATAG 360
 AAATCATCAT ATCTCCAGCT GGAAAAATGA TCTTTTGTG GGTAATGGTG ATAAAGCTGC 420

55

5		AACCAAGATA ATTAAGATAC ATCCAACCTG GGATAGCAGA AAACAATACA TGATGATCTC	480
		AAGGATAGTT ATCTGGATAT GCCCTACTAT AGCTGATCCT GATGGGAAAT TGGCTGTAGC	540
10		TTTAATTGAT CCTAACAAGA GTGTTAATGC CAGAACTGTT TTGAAAGGGC AAGGAAGCAT	600
		TAAAGATCCC ATATGTTTGG TTTTTTATCT AAATTGGTCC ATTCCAAAAG TTAACAACAC	660
		TTCAGAGAAT TGTGTTTCAGC TTCATTTATT ATGTGATCAA GTTTACAAGA AAGATGTTTC	720
15		TTTTGCTAGT GTCATGTATT CTTGGACAAN AGAATTCTGT GATTCACCAA GAGCAGATCT	780
		GGATAAAAGC TGCATGATAA TACCCATCAA TAGGGCTATT AGAGCCAAAT CGCAAGCCTT	840
		CATTGAAGCC TGCAAGTTAA TCATACCTAA AGGCAATTCT GAAAAGCAAA TTAGAAGACA	900
		ACTGTCAGAG CTAAGTGCTA ATTTAGAGAA ATCTGTTGAA GAAGAGGAGA ATGTTACTGA	960
20		TAACAAGATA GAGATATCAT TTGATAATGA AATCTAAATA TGTTCATT TAATAATAAA	1020
		TAATATATAT TGTTCATAAT ATTTTGAATG TTTAAGTAAA AAATAAAGCA AGATAAAAAA	1080
		CTATATATAT ATATATATAT AGAAGTATAA AATATATATG TATTTGTGTT TAAAAACAAA	1140
25		TCAAAAACCA AAAAAGAAAA AAGAAAAAAT AAACAAAAAA CAAAAACAAA AACAAAAACA	1200
		AACAAAAAGC AAAAATAGA AAAAAGTTGA AAAAACCAA AAAAATTTTT TTTGTAAATA	1260
		AATAAGGCTC CGGCCAGATT TGGCTAAGA CCTTTTATT TGTTCATTATA CATTTTATTT	1320
		GTTTTTGTG ATTTTTATTT TTATTATTTT TATATTTTTT ATATAGTTTG CTTATTTAAC	1380
30		ACTTATTTAG ACAAATTAAA TTTATTTGAT TACAATCATT CTGCCTTATT TAATTTAAAA	1440
		CACATTTGGT GSTATATCCA ATGAATTTAA TCATATACCG CTGAAGTCTA GAGGAGGTCT	1500
		TCTTCTAGTG ATGGTGCTT TACCAGAAGA CGTGGAAACC AAAGAATAAT CATTAGTGTC	1560
35		TTCAATATAT TTTGTCTTGT AAGACTTGTT TCTAACATAG CCTCTACACA TTGTGGCAAC	1620
		AATAGAGCAG AGGTAAGCAA GAGCAAATAC AAAGAGTATG AGCAATACTA CTCTGACTGT	1680
		ATCAAGAAG GATCCAAAGT GGCTTGCTAT AAAGTTAAAA GGGCTTTTAA CATAGTCCCA	1740
40		AAAGCTCCAA ACTGATGTGT CAGAATTATA TTGCTGTTCC TCGTGTGCAT GTTGGTCATT	1800
		TTGATCAATT ATGTTTTCTG GTTCCAGCAC AGCAACAGAA TCTACAAGTG CCTCAACTGA	1860
		GTATGATTGT TCTCCTTCTG GTTCTATAAT CATTTTTTGT TTTTCTGGGT TAGAAGTGCA	1920
		GAACATTGTC AAGTTATACT TATTAGCACC TTTCTTTACT GCTATCTGGT ATGTTGACAA	1980
45		TGAACATTGT TTCATGGTTA ACCTTGCAGA AAAAGTTATG TCTGATATAA ATGAGGCAGC	2040
		ACACCTCAGC CCTTGGCTAC ATAAGAAACA TCCCTTACAG CTAAAGAGA CAGAACTCAA	2100
		TATAGGCTTT TTTGGTACAG TTTTAAACAA TTCAGAAGGT AGATCCAAAA CAATTTTAAG	2160
50		CTTACCTAGA CTAAAGATCT TTTCCATATA AAAACTATTC TGGTCAGTAA ACTGAACTGG	2220
		AATGTCCGAT ATTTGGTTCA AACCTGTTTT AAATCTGTAT GTGTCATAAC CACATGATTT	2280
		TATCGTAATT GTTTTTTTAC CAATTGCTGA ACAATCCAG GACAGATCGT TTGTATCTAA	2340

5	TGTTTTCTTA GAGAAATGG GATCACCTTG GTGTGAAAGT TGAGGATGAC CAAACATTTT	2400
	TGATGGATTA TTTAATCTAG CTATGTTTCC CGCATATACG TGACTATCAG GTCCATGAGC	2460
10	TATCAGCTGG CCTATTGTTA AGCCATCATT ATGGAAATCC GCTAATATAT CAGCCTGGAA	2520
	ATATCCTGAT TCAGATGGGA CTTCTCAGA TACAGTGAAA CACTTTGCTC CCACAAATCC	2580
	AGATATACAT ACTTCAGACT TGATTGTTGA TTTAATAACA GAATAAATCC TGAAAGATTG	2640
	ATCCATATCA TACACATTC TACAAAACCC ACAAGTGGCT CCTTCATTGA TAGCCAAACA	2700
15	CCPAACCTCT TCACAACCCC AGTAAGATGT TGGTGTATG CAGAAATCTT GATACCCAGT	2760
	TATCGGTTGT TCTTTTCTGC AATCTGAGCA TTTACCTGTG CATGTTGAAA AGAAATCAGT	2820
	GTGGGTGCTT TGTATAGGAG CTGTAGTGTA TTGTTTCAGAA ACATCATACT GTATTCTAAC	2880
20	TTTTTTAATA TAAACAACAA ACTTCTGAGC AGTGCTAGAA CTTTGTGTCAT TAAGAGAGAA	2940
	AACTGTGCCC CCACCTGATA ATAAAGATTC TTCTATCATG TATCTATATT TTCCATCTAT	3000
	CACCGAGTCA AATATGAGAG ATTTTCTTGG AAAATGCTT TCAGGTATGT CTGATTCAAT	3060
25	AGATTTAAGT GCATCTCCAG AAATGTATCC ATATTTTCA GTTTTATTGT AGAAATCAAT	3120
	TATACCATTG CTAAGCCTTT TCATGAAGTG TAGATTCACA GCATTCAATC CCAATGTGTC	3180
	ACCAGAATAT TCTAAGAACC CATTATCTAA AGGCTTGCTT TGGAAAATAG AGGCATACTC	3240
	ACAACCAAAT CTGCATTGA CAAAAGTTAC TAAAGCATTT TCAGTTATCC TGCCTTTGCA	3300
30	TTCTTGATAA GGTATACAAT CCATAGGACC TTCTGTCACA ACATTGGTTA GAAAGTTAGA	3360
	TTCTACAATA GAATTTCTTT TAATAGCACA GAAGCATTGG TCTTTTTCAG GACATTTGTC	3420
	ATATCTGTTT GTAACAAAGC GGTACAAACC AGGGACATAA TAACAGCTAT CCAAACACTG	3480
35	AGCAGTTTGA GCCATAGACA TAGGCATCTG TGACAAAATC AGAAATCCTA TCAAAGTTTC	3540
	TGTGACTGCT TTTAGGAAAG AGAGGCCTAT TTTTGTATTA ACTATCAAAT GGAACCATTC	3600
	AATGCTAGCC CAGTTGTATT TTTTATTCTT CTCTGCTGTT CTAGTTATTA TAGGACATTC	3660
	TTCTGAGTGT TCTTCAGAGG CTTTGTTTTT GTTACAAATG CATAATTTTG AGCATTTCATG	3720
40	GGTTACCAAA CATAAATTTT CACAGACCTT ACATTTCAAG GGAAAATAAG ACCATAAATA	3780
	ATTTATCAGT AGTAGTATAG GATACGTTAT CAATCCCAGA AGATCATACC CATAGAACAG	3840
	TGTTTTAGAT GTTTTGTTTA CCAAGTACCT TATAGGGAAA TAGACAATCA GAGCAATCAT	3900
45	GATCAATCTA AACCATGAGA AGTTGATGCA AGCAGTTTGT TTGTAAATAT TTTTGGAGTA	3960
	CTTTATAATA CAATCTCTAA CTCTTTTGTG CACTAAAGGA ACTTTAGAAG ACTTGTCACC	4020
	GCACAATAGG TTATGCTTAC CATCCATATT TTCTTCTGTG AAAGTCAAAC TAACTGAGCC	4080
	AGAGAAGCTT ATTATGGAAT GGCTCATGTC ACTTCCTTCT CTTTGTGACGA CGTAACCCAT	4140
50	GATTTTCTCA GGTGTAGTTA ATGAACTGT ATAAGAATTA ACTATGTTTG TTTTGTATAT	4200
	TTTACAATCA CCTGAGAATT TCACACTCTG GAGAGAGACT GTGCCATTAG TTGGTCTAGA	4260

5

ATTGTACATG ATTGGATAAT TGTAATTCTC CAAACTTTCA ATTATATAGA ATTTAGTTCC 4320
TATAGATAAT TTCCTTTTGT TATCGATTTT TGTATTGGT ACAACTGGAA CAGTTTCAAA 4380
10 GCTTCTTGGC AATTCAGAAG ATCCTTCACA GTTTCCTCAAT TTAGTTATAG TGTCACTGAT 4440
ACATGAATAT ATAACACCAT TGCTTTCTAC TTGGTAATAA ACATTGAATG TTGAACTCC 4500
TTTAATGCTA CAAGTCAAAC TTGAAGCATT TAGGCATGGA TTTGGTAAAT CCATAACTGA 4560
TATAGTTGTT GGTGTAGAAG ACAATCCACT TGGAGATTGA GGTACCTCAT TATTGGCAAG 4620
15 AACAGTTTGA GTATCTCGTG TTGGTCTAAG GGTTTTACCT GTTGCAATTCT GGAGCATTTT 4680
AGCCAAAGTA TCTAGAATTT CATTTTTATG ATCTACAGAA CGGTCATAAT AAGCTTCATC 4740
ATAAATTTCT GGATGATCGC CCCTTTCAAC ATGAATCTTT GCATCTGTCT CCTTTAATGC 4800
20 CATAAAGGAT AAGATAACAG AAGTAACAAC TAGTGTACAT AACTAATTT TAACAAGTAA 4860
CTCGCACATC TTTAGAATTT TCATTCTAAA AAGTCGAATA AACTAGTTC TAAAATTGCT 4920
TTATGAGTTT GATCTGTTGT ATGTAGAGTT TTGTTTGCAC TGATTGCTCT 4970

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 912 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATGAATAGTT TTTTCAAATC ACTCAGATCA TCTAGCAGCA GGGAGCTAGA TCACCCTAGG 60
40 GTTACAACATA CCCTCTCTAA ACAAGGAGCA GACATTGTTG TACACAATCC TTCTGCTAAT 120
CACAACAACA AGGAAGTTCT CCAAAGAGCC ATGGATAGCT CTAAAGGGAA GATTTTGATG 180
AACAATACAG GCACCTCATC ACTAGGCACA TATGAGTCTG ACCAGATATC TGAATCAGAG 240
45 TCTTATGATC TTTCTGCTAG AATGATTGTT GATACAAATC ATCATATCTC CAGCTGGAAA 300
AATGATCTTT TTGTAGGTAA TGGTGATAAA GCTGCAACCA AGATAATTAA GATACATCCA 360
ACCTGGGATA GCAGAAAACA ATACATGATG ATCTCAAGGA TAGTTATCTG GATATGCCCT 420
50 ACTATAGCTG ATCCTGATGG GAAATTGGCT GTAGCTTTAA TTGATCCTAA CAAGAGTGTT 480
AATGCCAGAA CTGTTTTGAA AGGGCAAGGA AGCATTAAAG ATCCTATATG TTTTGTITTT 540
TATCTAAATT GGTCCATTCC AAAAGTTAAC AACACTTCAG AGAATTGTGT TCAGCTTCAT 600

55

5

TTATTATGTG ATCAAGTTTA CAAGAAAGAT GTTTCCTTTG CTAGTGTGAT GTATTCTTGG 660
 ACAAAGAAT TCTGTGATTC ACCAAGAGCA GATCTGGATA AAAGCTGCAT GATAATACCC 720
 ATCAATAGGG CTATTAGAGC CAAATCGCAA GCCTTCATTG AAGCCTGCAA GTTAATCATA 780
 CCTAAAGGCA ATTCTGAAAA GCAAATTAGA AGACAACTTG CAGAGCTAAG TGCTAATTTA 840
 GAGAAATCTG TTGAAGAAGA GGAGAATGTT ACTGATAACA AGATAGAGAT ATCATTGAT 900
 AATGAAATCT AA 912

15

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 473 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

20

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

30

ATATGTTTTT ATTAAATAAT AAATAATATA TATTGTTCAT AATATTTTGA ATGTTTAAGT 60
 AAAAAATAAA GCAAGATAAA AACTATATA TATATATATA TATAGAAGTA TAAATATAT 120
 ATGTATTTGT GTTAAAAAC AAATCAAAA CCAAAAAAGA AAAAAGAAAA AATAACAAA 180
 AAACAAAAAC AAAACAAAA ACAACAAAA AGCAAAAAAT AGAAAAAGT TGAAAAAAC 240
 CAAAAAATT TTTTTGTAA ATAAATAAGG CTCCGCCAG ATTTGGTCTA AGACCTTTTT 300
 ATTTGTTTTT ATACATTTTA TTGTTTTTG TTGATTTTAA TTTTATTAT TTTTATATT 360
 TTTATATAGT TTGCTTATT AACACTTATT TAGACAAATT AAATTTATT GATTACAATC 420
 ATTCTGCCTT ATTTAATTTA AAACACATT GGTGTATATT CCAATGAATT TAA 473

40

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3414 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

45

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

50

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

55

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

	TCATATACCG CTGAAGTCTA GAGGAGGTCT TCTTCTAGTG ATGGTGICTT TACCAGAAGA	60
10	CGTGGAACC AAAGAATAAT CATTAGTGTC TTCAATATAT TTEGTCTTGT AAGACTTGTT	120
	TCTAACATAG CCTCTACACA TTGTGGCAAC AATAGAGCAG AGGTAAGCAA GAGCAAATAC	180
	AAAGAGTAIG AGCAATACTA CTCTGACTGT ATCAAAGAAG GATCCAAAGT GGCTTGCTAT	240
15	AAAGTTAAAA GGGCTTTTAA CATAGTCCCA AAAGCTCCAA ACTGATGTGT CAGAATTATA	300
	TTGCTGTTCC TCGTGTGCAT GTTGGTCATT TTSATCAATT ATGTTTTCTG GTTCCAGCAC	360
	AGCAACAGAA TCTACAAGTG CCTCAACTGA GTATGATTTG TCTCCTTCTG GTTCTATAAT	420
20	CATTTTTTGT TTTTCTGGGT TAGAAGTGCA GAACATTGTC AAGTTATACT TATTAGCACC	480
	TTTCTTTACT GCTATCTGGT ATGTTGACAA TGAACATTGT TTCATGGTTA ACCTTGACAGA	540
	AAAAGTTATG TCTGATATAA ATGAGGCAGC ACACCTCAGC CCTTGGCTAC ATAAGAAACA	600
	TCCCTTACAG CTTAAAGAGA CAGAACTCAA TATAGGCTTT TTTGGTACAG TTTTAAACAA	660
25	TTCAGAAGGT AGATCCAAAA CAATTTTAAG CTTACCTAGA CTAAAGATCT TTTCCATATA	720
	AAAACATTTC TGGTCAGTAA ACTGAACTGG AATGTCCGAT ATTTGGTTCA AACCTGTTTT	780
	AAATCTGTAT GTGTCATAAC CACATGATTT TATCGTAATT GTTTTTTTTAC CAATTGCTGA	840
30	ACAATCCCAG GACAGATCGT TTGTATCTAA TGTTTTCTTA GAGAAAATGG GATCACCTTG	900
	GTGTGAAAGT TGAGGATGAC CAAACATTTT TGATGGATTA TTTAATCTAG CTATGTTTCC	960
	CGCATATACG TGACTATCAG GTCCATGAGC TATCAGCTGG CCTATTGTTA AGCCATCATT	1020
	ATGGAAAATCC GCTAATATAT CAGCCTGGAA ATATCCTGAT TCAGATGGGA CTTCTCAGAA	1080
35	TACAGTGAAA CACTTTGCTC CCACAAATCC AGATATACAT ACTTCAGACT TGATTGTTGA	1140
	TTTAATAACA GAATAAATCC TGAAAGATTG ATCCATATCA TACACATTC TACAAAACCC	1200
	ACAAGTGGCT CCTTCATTGA TAGCCAAACA CCAAACCTCT TCACAACCCC AGTAAGATGT	1260
40	TGGTGTTATG CAGAAATCTT GATACCCAGT TATCGGTTGT TCTTTTCTGC AATCTGAGCA	1320
	TTTACCTGTG CATGTTGAAA AGAAATCAGT GTGGGTGCTT TGTATAGGAG CTGTAGTGTA	1380
	TTGTTACAGAA ACATCATACT GTATTCTAAC TTTTTTAATA TAAACAACAA ACTTCTGAGC	1440
45	AGTGCTAGAA CTTTTGTCAT TAAGAGAGAA AACTGTGCCC CCACCTGATA ATAAAGATTC	1500
	TTCTATCATG TATCTATATT TTCCATCTAT CACCGAGTCA AATATGAGAG ATTTTCTTGG	1560
	AAAAATGCTT TCAGGTATGT CTGATTCATT AGATTTAAGT GCATCTCCAG AAATGTATCC	1620
	ATATTTTTCA GTTTTATTGT AGAAATCAAT TATACCATT CTAAGCCTTT TCATGAAGTG	1680
50	TAGATTCACA GCATTCAATC CCAATGTGTC ACCAGAATAT TCTAAGAACC CATTATCTAA	1740
	AGGCTTGCTT TGGAAAATAG AGGCATACTC ACAACCAAA CTGCATTTGA CAAAAGTTAC	1800
	TAAAGCATTT TCAGTTATCC TGCCTTTGCA TTCTTGATAA GGTATACAAT CCATAGGACC	1860

55

5

10

15

20

25

30

35

40

45

TTCTGTCACA ACATTGGTTA GAAAGTTAGA TTCTACAATA GAATTTTCTT TAATAGCACA 1920
 GAAGCATTGG TCTTTTCAG GACATTTGTC ATATCTGTTT GTAACAAAGC GGTACAAACC 1980
 AGGGACATAA TAACAGCTAT CCAAACACTG AGCAGTTTGA GCCATAGACA TAGGCATCTG 2040
 TGACAAAATC AGAAATCCTA TCAAAGTTTC TGTGACTGCT TTTAGGAAAG AGAGGCCTAT 2100
 TTTTGTATTA ACTATCAAAT GGAACCATTC AATGCTAGCC CAGTTGTATT TTTTATTCTT 2160
 CTCTGCTGTT CTAGTTATTA TAGGACATTC TTCTGAGTGT TCTTCAGAGG CTTTGTTTTT 2220
 GTTACAAATG CATAATTTTG AGCATTCATG GGTACCAA CATAAATTTT CACAGACCTT 2280
 ACATTTCAGG GGAAATAAG ACCATAAATA ATTTATCAGT AGTAGTATAG GATACGTTAT 2340
 CAATCCCAGA AGATCATACC CATAGAACAG TGTTTGTAGT GTTTTGTTTA CCAAGTACCT 2400
 TATAGGGAAA TAGACAATCA GAGCAATCAT GATCAATCTA AACCATGAGA AGTTGATGCA 2460
 AGCAGTTTGT TTGTAAATAT TTTTGGAGTA CTTTATAATA CAATCTCTAA CTCTTTTGTC 2520
 CACTAAAGGA ACTTTAGAAG ACTTGTACCC GCACAAATAGG TTATGCTTAC CATCCATATT 2580
 TTCTTCTGTG AAAGTCAAAC TAACTGAGCC AGAGAAGCTT ATTATGGAAT GGCTCATGTC 2640
 ACTTCCTTCT CTTTGTACGA CGTAACCCAT GATTTTCTCA GGTGTAGTTA ATGAACTGT 2700
 ATAAGAATTA ACTATGTTTG TTTTGTATAT TTTACAATCA CCTGAGAATT TCACACTCTG 2760
 GAGAGAGACT GTGCCATTAG TTGGTCTAGA ATTGTACATG ATTGGATAAT TGTAATTCTC 2820
 CAAACTTTCA ATTATATAGA ATTTAGTTCC TATAGATAAT TTCCTTTTGT TATCGATTTT 2880
 TGTATTGGT ACAACTGGAA CAGTTTCAAA GCTTCTTGGC AATTCAGAAG ATCCTTCACA 2940
 GTTTCCCAAT TTAGTTATAG TGTCACATGAT ACATGAATAT ATAACACCAT TGCTTTCTAC 3000
 TTGGTAATAA ACATTGAATG TTGAACTCC TTTAATGCTA CAAGTCAAAC TTGAAGCATT 3060
 TAGGCATGGA TTTGGTAAAT CCATAACTGA TATAGTTGTT GGTGTAGAAG ACAATCCACT 3120
 TGGAGATTGA GGTACCTCAT TATTGGCAAG AACAGTTTGA GTATCTCGTG TTGGTCTAAG 3180
 GGTTTTACCT GTTGCAATCT GGAGCATTTT AGCCAAAGTA TCTAGAATTT CATTTTTATG 3240
 ATCTACAGAA CGGTCATAAT AAGCTTCATC ATAAATTTCT GGATGATCGC CCCTTTCAAC 3300
 ATGAATCTTT GCATCTGTCT CCTTTAATGC CATAAAGGAT AAGATAACAG AAGTAACAAC 3360
 TAGTGTACAT ACACTAATTT TAACAAGTAA CTCGCACATC TTTAGAATTT TCAT 3414

(2) INFORMATION FOR SEQ ID NO:18:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

50

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

55

5

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AGAGCAATCA GTGCATCAAA ATTATATCTA GCCGAA

36

(2) INFORMATION FOR SEQ ID NO:19:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

20

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTGTTGTATG TAGAGTTTGT TTTGCACTGA TTGCTC

36

30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4970 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

35

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

45

AGAGCAATCA GTGCAAACAA AACTCTACAT ACAACAGATC AAACTCATAA AGCAATTTTA 60

GAACTAGTGT TATTCGACTT TTTAGAATGA AAATTCTAAA GATGTGCGAG TTAATTGTGA 120

AAATTAGTGT ATGTACACTA GTTGTTACTT CTGTTATCTT ATCCTTTATG GCATTAAAGG 180

50

AGACAGATGC AAAGATTCAT GTTGAAAGGG GCGATCATCC AGAAATTTAT GATGAAGCTT 240

ATTATGACCG TTCTGTAGAT CATAAAATG AAATTCTAGA TACTTTGGCT GAAATGCTCC 300

AGAATGCAAC AGGTAAAACC CTTAGACCAA CACGAGATAC TCAAACGTCT CTTGCCAATA 360

55

5

	ATGAGGTACC TCAATCTCCA AGTGGATTGT CTTCTACACC AACAACTATA TCAGTTATGG	420
	ATTTACCAAAA TCCATGCCTA AATGCTTCAA GTTTGACTTG TAGCATTAAA GGAGTTTCAA	480
10	CATTCAATGT TTATTACCAA GTAGAAAGCA ATGGTGTAT ATATTCAATGT ATCAGTGACA	540
	CTATAACTAA ATTGGGAAAC TGTGAAGGAT CTTCTGAATT GCCAAGAAGC TTTGAAACTG	600
	TTCCAGTTGT ACCAATAACA AAAATCGATA ACAAAGGAA ATTATCTATA GGAACATAAT	660
	TCTATATAAT TGAAAGTTTG GAGAATTACA ATTATCCAAT CATGTACAAT TCTAGACCAA	720
15	CTAATGGCAC AGTCTCTCTC CAGAGTGTGA AATTCTCAGG TGATTGTAAA ATATCAAAAA	780
	CAAACATAGT TAATTCTTAT ACAGTTTCAT TAACTACACC TGAGAAAATC ATGGGTACG	840
	TCGTCAAAAG AGAAGGAAGT GACATGAGCC ATTCCATAAT AAGCTTCTCT GGCTCAGTTA	900
20	GTTTGACTTT CACAGAAGAA AATATGGATG GTAAGCATAA CCTATTGTGC GGTGACAAGT	960
	CTTCTAAAGT TCCTTTAGTG GACAAAAGAG TTAGAGATTG TATTATAAAG TACTCCAAAA	1020
	ATATTTACAA ACAAACGTCT TGCATCAACT TCTCATGGTT TAGATTGATC ATGATTGCTC	1080
25	TGATTGTCTA TTTCCCTATA AGGTACTTGG TAAACAAAAC ATCTAAAACA CTGTTCTATG	1140
	GGTATGATCT TCTGGGATTG ATAACGTATC CTATACTACT ACTGATAAAT TATTTATGGT	1200
	CTTATTTTCC CTTGAAATGT AAGGTCTGTG GAAATTTATG TTTGGTAACC CATGAATGCT	1260
	CAAAATTATG CATTGTGAAC AAAAACAAAG CCTCTGAAGA ACACTCAGAA GAATGTCCTA	1320
30	TAATAACTAG AACAGCAGAG AAGAATAAAA AATACAACGT GGCTAGCATT GAATGGTTCC	1380
	ATTTGATAGT TAATACAAAA ATAGGCCTCT CTTTCCTAAA AGCAGTCACA GAACTTTTGA	1440
	TAGGATTTCT GATTTTGTCA CAGATGCCTA TGTCTATGGC TCAAACGTCT CAGTGTGTTG	1500
35	ATAGCTGTTA TTATGTCCCT GTTTGTGACC GCTTTGTTAC AAACAGATAT GACAAATGTC	1560
	CTGAAAAAGA CCAATGCTTC TGTGCTATTA AAGAAAATTC TATTGTAGAA TCTAACTTTC	1620
	TAACCAATGT TGTGACAGAA GGTCCCTATGG ATTGTATACC TTATCAAGAA TGCAAAGGCA	1680
40	GGATAACTGA AAATGCTTTA GTAACTTTGT TCAAATGCAG ATTTGGTTGT GAGTATGCCT	1740
	CTATTTTCCA AAGCAAGCCT TTAGATAATG GGTTCCTAGA ATATTCTGGT GACACATTGG	1800
	GATTGAATGC TGTGAATCTA CACTTCATGA AAAGGCTTAG GAATGGTATA ATTGATTCT	1860
	ACAATAAAAC TGAAAAATAT GGATACATTT CTGGAGATGC ACTTAAATCT AATGAATCAG	1920
45	ACATACCTGA AAGCATTTTT CCAAGAAAAT CTCTCATATT TGAATCGGTG ATAGATGGAA	1980
	AATATAGATA CATGATAGAA GAATCTTTAT TATCAGGTGG GGGCACAGTT TTCTCTCTTA	2040
	ATGACAAAAG TTCTAGCACT GCTCAGAAGT TTGTTGTTTA TATTAATAAA GTTAGAATAC	2100
50	AGTATGATGT TTCTGAACAA TAACTACAG CTCCTATACA AAGCACCAC ACTGATTCT	2160
	TTTCAACATG CACAGGTAAA TGCTCAGATT GCAGAAAAGA ACAACCGATA ACTGGGTATC	2220
	AAGATTTCTG CATAACACCA ACATCTTACT GGGGTTGTGA AGAGGTTTGG TGTGTTGGCTA	2280

55

5	TCAATGAAGG AGCCACTTGT GGGTTTTGTA GAAATGTGTA TGATATGGAT CAATCTTTCA	2340
	GGATTTATTC TGTATTAAA TCAACAATCA AGTCTGAAGT ATGTATATCT GGATTTGTGG	2400
10	GAGCAAAGTG TTTCAGTGA TCTGAGGAAG TCCCCTCTGA ATCAGGATAT TTCCAGGCTG	2460
	ATATATTAGC GGATTTCCAT AATGATGGCT TAACAATAGG CCAGCTGATA GCTCATGGAC	2520
	CTGATAGTCA CGTATATGCG GGAAACATAG CTAGATTAAA TAATCCATCA AAAATGTTTG	2580
15	GTCATCCTCA ACTTTCACAC CAAGGTGATC CCATTTTCTC TAAGAAAACA TTAGATACAA	2640
	ACGATCTGTC CTGGGATTGT TCAGCAATTG GTAAAAAAC AATTACGATA AAATCATGTG	2700
	GTTATGACAC ATACAGATTT AAAACAGGTT TGAACCAAA ATCGGACATT CCAGTTCAGT	2760
	TTACTGACCA GAATAGTTTT TATATGGAAA AGATCCTTAG TCTAGGTAAG CTTAAAATTG	2820
20	TTTTGGATCT ACCTTCTGAA TTGTTTAAAA CTGTACCAA AAAGCCTATA TTGAGTTCTG	2880
	TCTCTTTAAG CTGTAAGGGA TGTTCCTTAT GTAGCCAAGG GCTGAGGTGT GCTGCCTCAT	2940
	TTATATCAGA CATAACTTTT TCTGCAAGGT TAACCATGAA ACAATGTTC TGTCAACAT	3000
25	ACCAGATAGC AGTAAAGAAA GGTGCTAATA AGTATAACTT GACAATGTTC TGCACCTCTA	3060
	ACCCAGAAAA ACAAAAATG ATTATAGAAC CAGAAGGAGA CAAATCATAC TCAGTTGAGG	3120
	CACTTGTAGA TTCTGTTGCT GTGCTGGAAC CAGAAAACAT AATTGATCAA AATGACCAAC	3180
30	ATGCACACGA GGAACAGCAA TATAATTCTG ACACATCAGT TTGGAGCTTT TGGGACTATG	3240
	TTAAAAGCCC TTTTAACTTT ATAGCAAGCC ACTTTGGATC CTTCTTTGAT ACAGTCAGAG	3300
	TAGTATTGCT CATACTCTTT GTATTGCTC TTGCTTACCT CTGCTCTATT GTTGCCACAA	3360
	TGTGTAGAGG CTATGTTAGA AACAAGTCTT ACAAGACAAA ATATATTGAA GACACTAATG	3420
35	ATTATTCTTT GGTTCACAG TCTTCTGGTA AAGACCCAT CACTAGAAGA AGACCTCCTC	3480
	TAGACTTCAG CGGTATATGA TTAAATTCAT TGGATATAC ACCAAATGTG TTTTAAATTA	3540
	AATAAGGCAG AATGATTGTA ATCAAATAAA TTTAATTTGT CTAAATAAGT GTTAAATAAG	3600
40	CAAACTATAT AAAAAATATA AAAATAATAA AAATAAAAA CAACAAAAAC AAATAAAATG	3660
	TATAAAAACA AATAAAAAGG TCTTAGACCA AATCTGGCCG GAGCCTTATT TATTTACAAA	3720
	AAAAATTTTT TTGGTTTTTT TCAACTTTTT TCTATTTTT GCTTTTTGTT TGTTTTTGTT	3780
	TTTGTTTTTG TTTTTTGTTT ATTTTTCTT TTTCTTTTT TGTTTTTTGA TTTGTTTTTA	3840
45	AACACAAATA CATATATATT TTATACTTCT ATATATATAT ATATATATAG TTTTTATCT	3900
	TGCTTTATTT TTTACTTAAA CATTCAAAAT ATTATGAACA ATATATATTA TTTATTATTA	3960
	AATGAAAACA TATTTAGATT TCATTATCAA ATGATATCTC TATCTTGTTA TCAGTAACAT	4020
50	TCTCCTCTTC TTCAACAGAT TTCTCTAAAT TAGCACTTAG CTCTGCAAGT TGTCTTCTAA	4080
	TTTGCTTTTC AGAATTGCCT TTAGGTATGA TTAAGTTGCA GGCTTCAATG AAGGCTTGCG	4140
	ATTTGGCTCT AATAGCCCTA TTGATGGGTA TTAATCATGA GCTTTTATCC AGATCTGCTC	4200

5

TTGGTGAATC ACAGAATTCT TTTGTCCAAG AATACATGAC ACTAGCAAAA GAAACATCTT 4260
 TCTTGTAAC TTGATCACAT AATAAATGAA GCTGAACACA ATTCTCTGAA GTGTTGTAA 4320
 10 CTTTGGGAAT GGACCAATTT AGATAAAAAA CAAAACATAT AGGATCTTTA ATGCTTCCTT 4380
 GCCCTTTCAA AACAGTTCTG GCATTAACAC TCTTGTTAGG ATCAATTAAA GCTACAGCCA 4440
 ATTTCCCATC AGGATCAGCT ATAGTAGGGC ATATCCAGAT AACTATCCTT GAGATCATCA 4500
 15 TGTATTGTTT TCTGCTATCC CAGGTTGGAT GTATCTTAAT TATCTTGGTT GCAGCTTTAT 4560
 CACCATTACC TACAAAAGA TCATTTTCC AGCTGGAGAT ATGATGATTT GTATCAACAA 4620
 TCATTCTAGC AGAAAGATCA TAAGACTCTG ATTCAGATAT CTGGTCAGAC TCATATGTGC 4680
 CTAGTGATGA GGTGCCTGTA TTGTTTCATCA AAATCTTCCC TTAGAGCTA TCCATGGCTC 4740
 20 TTTGGAGAAC TTCCTTGTTG TTGTGATTAG CAGAAGGATT GTGTACAACA ATGTCTGCTC 4800
 CTTGTTTAGA GAGGGTAGTT GTAACCCTAG GTGATCTAG CTCCCTGCTG CTAGATGATC 4860
 TGAGTGATTT GAAAAAATA TTCATCTCTG GACGGTAGAT TAAGAATTAA AATATTGAGA 4920
 25 AGATAATGAT TGAATTCGGC TAGATATAAT TTTGATGCAC TGATTGCTCT 4970

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3414 base pairs
 (B) TYPE: nucleic acid
 30 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown
 (ii) MOLECULE TYPE: cDNA
 (iii) HYPOTHETICAL: NO
 35 (iii) ANTI-SENSE: YES
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATGAAAATTC TAAAGATGTG CGAGTTACTT GTTAAAATTA GTGTATGTAC ACTAGTTGTT 60
 ACTTCTGTTA TCTTATCCTT TATGGCATT AAGGAGACAG ATGCAAAGAT TCATGTTGAA 120
 AGGGGCGATC ATCCAGAAAT TTATGATGAA GCTTATTATG ACCGTTCTGT AGATCATAAA 180
 45 AATGAAATTC TAGATACTTT GGCTGAAATG CTCCAGAATG CAACAGGTAA AACCCTTAGA 240
 CCAACACGAG ATACTCAAAC TGTTCCTGCC AATAATGAGG TACCTCAATC TCCAAGTGGA 300
 TTGTCTTCTA CACCAACAAC TATATCAGTT ATGGATTTAC CAAATCCATG CCTAAATGCT 360
 50 TCAAGTTTGA CTTGTAGCAT TAAAGGAGTT TCAACATTCA ATGTTTATTA CCAAGTAGAA 420
 AGCAATGGTG TTATATATTC ATGTATCAGT GACACTATAA CTAAATTGGG AAAGTGTGAA 480
 GGATCTTCTG AATTGCCAAG AAGCTTTGAA ACTGTTCCAG TTGTACCAAT AACAAAAATC 540

55

5

	GATAACAAAA	GGAAATTATC	TATAGGAACT	AAATTCTATA	TAATTGAAAG	TTTGGAGAAT	600
	TACAATTATC	CAATCATGTA	CAATCTAGA	CCAACTAATG	GCACAGTCTC	TCTCCAGAGT	660
10	GTGAAATTCT	CAGGTGATTG	TAAAATATCA	AAAACAAACA	TAGTTAATTC	TTATACAGTT	720
	TCATTAACCTA	CACCTGAGAA	AATCATGGGT	TACGTCGTCA	AAAGAGAAGG	AAGTGACATG	780
	AGCCATTCCA	TAATAAGCTT	CTCTGGCTCA	GTTAGTTTGA	CTTTCACAGA	AGAAAATATG	840
	GATGGTAAGC	ATAACCTATT	GTGCGGTGAC	AAGTCTTCTA	AAGTTCCTTT	AGTGGACAAA	900
15	AGAGTTAGAG	ATTGTATTAT	AAAGTACTCC	AAAAATATTT	ACAAACAAAC	TGCTTGCAATC	960
	AACTTCTCAT	GGTTTAGATT	GATCATGATT	GCTCTGATTG	TCTATTTCCC	TATAAGGTAC	1020
	TTGGTAACAA	AAACATCTAA	AACACTGTTT	TATGGGTATG	ATCTTCTGGG	ATTGATAACG	1080
20	TATCCTATAC	TACTACTGAT	AAATTATTTA	TGGTCTTATT	TTCCCTTGAA	ATGTAAGGTC	1140
	TGTGGAAATT	TATGTTTGGT	AACCCATGAA	TGCTCAAAAT	TATGCATTTG	TAACAAAAAC	1200
	AAAGCCTCTG	AAGAACACTC	AGAAGAATGT	CCTATAATTA	CTAGAACAGC	AGAGAAGAAT	1260
25	AAAAAATACA	ACTGGGCTAG	CATTGAATGG	TTCCATTTGA	TAGTTAATAC	AAAAATAGGC	1320
	CTCTCTTTCC	TAAAAGCAGT	CACAGAACT	TTGATAGGAT	TTCTGATTTT	GTCACAGATG	1380
	CCTATGTCTA	TGGCTCAAAC	TGCTCAGTGT	TTGGATAGCT	GTTATTATGT	CCCTGGTTGT	1440
	GACCGCTTTG	TTACAAACAG	ATATGACAAA	TGTCCTGAAA	AAGACCAATG	CTTCTGTGCT	1500
30	ATTAAAGAAA	ATTCTATTGT	AGAATCTAAC	TTTCTAACCA	ATGTTGTGAC	AGAAGGTCCT	1560
	ATGGATTGTA	TACCTTATCA	AGAATGCAAA	GGCAGGATAA	CTGAAAATGC	TTTAGTAACT	1620
	TTTGTCAAAAT	GCAGATTTGG	TTGTGAGTAT	GCCTCTATTT	TCCAAAGCAA	GCCTTTAGAT	1680
35	AATGGGTTCT	TAGAATATTC	TGGTGACACA	TTGGGATTGA	ATGCTGTGAA	TCTACACTTC	1740
	ATGAAAAGGC	TTAGGAATGG	TATAATTGAT	TTCTACAATA	AAACTGAAAA	ATATGGATAC	1800
	ATTTCTGGAG	ATGCACTTAA	ATCTAATGAA	TCAGACATAC	CTGAAAGCAT	TTTTCCAAGA	1860
40	AAATCTCTCA	TATTTGACTC	GGTGATAGAT	GGAAAATATA	GATACATGAT	AGAAGAATCT	1920
	TTATTATCAG	GTGGGGGCAC	AGTTTTCTCT	CTTAATGACA	AAAGTTCTAG	CACTGCTCAG	1980
	AAGTTTGTG	TTTATATTAA	AAAAGTTAGA	ATACAGTATG	ATGTTTCTGA	ACAATACACT	2040
	ACAGCTCCTA	TACAAAGCAC	CCCACTGAT	TTCTTTTCAA	CATGCACAGG	TAAATGCTCA	2100
45	GATTGCAGAA	AAGAACAACC	GATAACTGGG	TATCAAGATT	TCTGCATAAC	ACCAACATCT	2160
	TACTGGGGTT	GTGAAGAGGT	TTGGTGTGTT	GCTATCAATG	AAGGAGCCAC	TTGTGGGTTT	2220
	TGTAGAAATG	TGTATGATAT	GGATCAATCT	TTCAGGATTT	ATTCTGTTAT	TAAATCAACA	2280
50	ATCAAGTCTG	AAGTATGTAT	ATCTGGATTT	GTGGGAGCAA	AGTGTTCAC	TGTATCTGAG	2340
	GAAGTCCCAT	CTGAATCAGG	ATATTCCAG	GCTGATATAT	TAGCGGATTT	CCATAATGAT	2400
	GGCTTAACAA	TAGGCCAGCT	GATAGCTCAT	GGACCTGATA	GTCACGTATA	TGCGGGAAAC	2460

55

5

10

15

20

25

30

35

40

45

50

ATAGCTAGAT TAAATAATCC ATCAAAAATG TTTGGTCATC CTCAACTTTC ACACCAAGGT 2520
 GATCCCATT TCTCTAAGAA AACATTAGAT ACAAACGATC TGTCTGGGA TTGTTAGCA 2580
 ATTGGTAAAA AAACAATTAC GATAAAATCA TGTGGTTATG ACACATACAG ATTTAAAACA 2640
 GGTTTGAACC AAATATCGGA CATTCCAGTT CAGTTTACTG ACCAGAATAG TTTTATATG 2700
 GAAAAGATCT TTAGTCTAGG TAAGCTTAAA ATTGTTTGG ATCTACCTTC TGAATTGTTT 2760
 AAAACTGTAC CAAAAAGCC TATATTGAGT TCTGTCTCTT TAAGCTGTAA GGGATGTTTC 2820
 TTAIGTAGCC AAGGGCTGAG GTGTGCTGCC TCATTATAT CAGACATAAC TTTTCTGCA 2880
 AGGTTAACCA TGAAACAATG TCATTGTCA ACATACCAGA TAGCAGTAAA GAAAGGTGCT 2940
 AATAAGTATA ACTTGACAAT GTTCTGCACT TCTAACCAG AAAACAAAA AATGATTATA 3000
 GAACCAGAAG GAGACAAATC ATACTCAGTT GAGGCACTTG TAGATTCTGT TGCTGTGCTG 3060
 GAACCAGAAA ACATAATTGA TCAAAATGAC CAACATGCAC ACGAGGAACA GCAATATAAT 3120
 TCTGACACAT CAGTTTGGAG CTTTGGGAC TATGTTAAA GCCCTTTTAA CTTTATAGCA 3180
 AGCCACTTTG GATCCTTCTT TGATACAGTC AGAGTAGTAT TGCTCATACT CTTTGTATTT 3240
 GCTCTTGCTT ACCTCTGCTC TATTGTTGCC ACAATGTGTA GAGGCTATGT TAGAAACAAG 3300
 TCTTACAAGA CAAATATAT TGAAGACACT AATGATTATT CTTTGGTTTC CACGTCTTCT 3360
 GGTAAAGACA CCATCACTAG AAGAAGACCT CCTCTAGACT TCAGCGGTAT ATGA 3414

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 912 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTAGATTTCA TTATCAAATG ATATCTCTAT CTTGTTATCA GTAACATTCT CCTCTTCTTC 60
 AACAGATTTT TCTAAATTAG CACTTAGCTC TGCAAGTTGT CTTCTAATTT GCTTTTCAGA 120
 ATTGCCTTTA GGTATGATTA ACTTGCAGGC TTCAATGAAG GCTTGCGATT TGGCTCTAAT 180
 AGCCCTATTG ATGGGTATTA TCATGCAGCT TTTATCCAGA TCTGCTCTTG GTGAATCACA 240
 GAATTCTTTT GTCCAAGAAT ACATGACACT AGCAAAAGAA ACATCTTTCT TGTAACCTTG 300
 ATCACATAAT AAATGAAGCT GAACACAATT CTCTGAAGTG TTGTTAACTT TTGGAATGGA 360

55

5

CCAATTTAGA TAAAAACAA AACATATAGC ATCTTTAATG CTTCCCTTGCC CTTTCAAAAC 420
 AGTTCTGGCA TTAACACTCT TGTTAGGATC AATTAAAGCT ACAGCCAATT TCCCATCAGG 480
 10 ATCAGCTATA GTAGGGCATA TCCAGATAAC TATCCTTGAG ATCATCATGT ATTGTTTTCT 540
 GCTATCCCAG GTTGGATGTA TCTTAATTAT CTTGGTTGCA GCTTTATCAC CATTACCTAC 600
 AAAAAGATCA TTTTCCAGC TGGAGATATG ATGATTGTGA TCAACAATCA TTCTAGCAGA 660
 AAGATCATAA GACTCTGATT CAGATATCTG GTCAGACTCA TATGTGCCTA GTGATGAGGT 720
 15 GCCTGTATTG TTCAACAAA TCTTCCCTTT AGAGCTATCC ATGGCTCTTT GGAGAACTTC 780
 CTTGTTGTTG TGATTAGCAG AAGGATTGTG TACAACAATG TCTGCTCCTT GTTTAGAGAG 840
 GGTAGTTGTA ACCCTAGGGT GATCTAGCTC CCTGCTGCTA GATGATCTGA GTGATTGAA 900
 20 AAAACTATTC AT 912

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 446 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Tobacco mosaic virus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGATCCGGAA CATGGTGGAG CACGACACGC TTGTCTACTC CAAAAATATC AAAGATACAG 60
 TCTCAGAAGA CCAAAGGGCA ATTGAGACTT TTCAACAAAG TTATTGTGAA GATAGTGGAA 120
 40 AAGGAAGGTG GCTCCTACAA ATGCCATCAT TGCGATAAAG GAAAGGCCAT CGTTGAAGAT 180
 GCCTCTGCCG ACAGTGGTCC CAAAGATGGA CCCCCACCCA CGAGGAGCAT CGTGGAAAAA 240
 GAAGACGTTT CAACCACGTC TTCAAAGCAA GTGGATTGAT GTGATATCTC CACTGACGTA 300
 AGGGATGACG CACAATCCCA CTATCCTTCG CAAGACCCTT CCTCTATATA AGGAAGTTCA 360
 45 TTTCATTTGG AGAGGACTTT TTACAACAAT TACCAACAAC AACAAACAAC AAACAACATT 420
 ACAATTACTA TTTACAATTA CCCGGG 446

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 303 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

55

5

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

10

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

15

Met Asn Ser Phe Phe Lys Ser Leu Arg Ser Ser Ser Ser Arg Glu Leu
1 5 10 15Asp His Pro Arg Val Thr Thr Thr Leu Ser Lys Gln Gly Ala Asp Ile
20 25 30

20

Val Val His Asn Pro Ser Ala Asn His Asn Asn Lys Glu Val Leu Gln
35 40 45Arg Ala Met Asp Ser Ser Lys Gly Lys Ile Leu Met Asn Asn Thr Gly
50 55 60

25

Thr Ser Ser Leu Gly Thr Tyr Glu Ser Asp Gln Ile Ser Glu Ser Glu
65 70 75 80Ser Tyr Asp Leu Ser Ala Arg Met Ile Val Asp Thr Asn His His Ile
85 90 95

30

Ser Ser Trp Lys Asn Asp Leu Phe Val Gly Asn Gly Asp Lys Ala Ala
100 105 110Thr Lys Ile Ile Lys Ile His Pro Thr Trp Asp Ser Arg Lys Gln Tyr
115 120 125Met Met Ile Ser Arg Ile Val Ile Trp Ile Cys Pro Thr Ile Ala Asp
130 135 140

35

Pro Asp Gly Lys Leu Ala Val Ala Leu Ile Asp Pro Asn Lys Ser Val
145 150 155 160Asn Ala Arg Thr Val Leu Lys Gly Gln Gly Ser Ile Lys Asp Pro Ile
165 170 175

40

Cys Phe Val Phe Tyr Leu Asn Trp Ser Ile Pro Lys Val Asn Asn Thr
180 185 190Ser Glu Asn Cys Val Gln Leu His Leu Leu Cys Asp Gln Val Tyr Lys
195 200 205

45

Lys Asp Val Ser Phe Ala Ser Val Met Tyr Ser Trp Thr Lys Glu Phe
210 215 220Cys Asp Ser Pro Arg Ala Asp Leu Asp Lys Ser Cys Met Ile Ile Pro
225 230 235 240Ile Asn Arg Ala Ile Arg Ala Lys Ser Gln Ala Phe Ile Glu Ala Cys
245 250 255

50

Lys Leu Ile Ile Pro Lys Gly Asn Ser Glu Lys Gln Ile Arg Arg Gln
260 265 270

Leu Ala Glu Leu Ser Ala Asn Leu Glu Lys Ser Val Glu Glu Glu

55

5

275

280

285

Asn Val Thr Asp Asn Lys Ile Glu Ile Ser Phe Asp Asn Glu Ile
 290 295 300

10

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 262 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

15

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

20

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

25

Met Asn Lys Ala Lys Ile Thr Lys Glu Asn Ile Val Lys Leu Leu Thr
 1 5 10 15

Gln Ser Asp Ser Leu Glu Phe Glu Glu Thr Gln Asn Glu Gly Ser Phe
 20 25 30

30

Asn Phe Thr Asp Phe Phe Thr Asn Asn Arg Glu Lys Ile Gln Asn Met
 35 40 45

Thr Thr Ala Ser Cys Leu Ser Phe Leu Lys Asn Arg Gln Ser Ile Met
 50 55 60

35

Arg Val Ile Lys Ser Ala Asp Phe Thr Phe Gly Ser Val Thr Ile Lys
 65 70 75 80

Lys Thr Arg Asn Asn Ser Glu Arg Val Gly Val Asn Asp Met Thr Phe
 85 90 95

Arg Arg Leu Asp Ala Met Val Arg Val His Leu Val Gly Met Ile Lys
 100 105 110

40

Asp Asn Gly Ser Ala Leu Thr Glu Ala Ile Asn Ser Leu Pro Ser His
 115 120 125

Pro Leu Ile Ala Ser Tyr Gly Leu Ala Thr Thr Asp Leu Lys Ser Cys
 130 135 140

45

Val Leu Gly Val Leu Leu Gly Gly Ser Leu Pro Leu Ile Ala Ser Val
 145 150 155 160

Leu Asn Phe Glu Ile Ala Ala Leu Val Pro Ala Ile Tyr Gln Asp Ala
 165 170 175

50

Lys His Val Glu Leu Gly Ile Asp Met Ser Lys Phe Ser Thr Lys Glu
 180 185 190

Ala Val Gly Lys Val Cys Thr Val Leu Lys Ser Lys Gly Tyr Ser Met
 195 200 205

55

5

Asn Ser Val Glu Ile Gly Lys Ala Lys Gln Tyr Ala Asp Ile Leu Lys
210 215 220

10

Ala Cys Ser Pro Lys Ala Lys Gly Leu Ala Ala Met Asp His Tyr Lys
225 230 235 240

Glu Gly Leu Thr Ser Ile Tyr Ser Met Phe Asn Ala Thr Ile Asp Phe
245 250 255

Gly Lys Asn Asp Ser Ile
260

15

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 449 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

20

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

25

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

30

Met Ser Ser Ala Met Tyr Glu Thr Ile Ile Lys Ser Lys Ser Ser Ile
1 5 10 15

Trp Gly Thr Thr Ser Ser Gly Lys Ala Val Val Asp Ser Tyr Trp Ile
20 25 30

35

His Asp Gln Ser Ser Gly Lys Lys Leu Val Glu Ala Gln Leu Tyr Ser
35 40 45

Asp Ser Arg Ser Lys Thr Ser Phe Cys Tyr Thr Gly Lys Val Gly Phe
50 55 60

40

Leu Pro Thr Glu Glu Lys Glu Ile Ile Val Arg Cys Phe Val Pro Ile
65 70 75 80

Phe Asp Asp Ile Asp Leu Asn Phe Ser Phe Ser Gly Asn Val Val Glu
85 90 95

45

Ile Leu Val Arg Ser Asn Thr Thr Asn Thr Asn Gly Val Lys His Gln
100 105 110

Gly His Leu Lys Val Leu Ser Ser Gln Leu Leu Arg Met Leu Glu Glu
115 120 125

Gln Ile Ala Val Pro Glu Ile Thr Ser Arg Phe Gly Leu Lys Glu Ser
130 135 140

50

Asp Ile Phe Pro Pro Asn Asn Phe Ile Glu Ala Ala Asn Lys Gly Ser
145 150 155 160

Leu Ser Cys Val Lys Glu Val Leu Phe Asp Val Lys Tyr Ser Asn Asn
165 170 175

55

5

Gln Ser Met Gly Lys Val Ser Val Leu Ser Pro Thr Arg Ser Val His
 180 185 190
 Glu Trp Leu Tyr Thr Leu Lys Pro Val Phe Asn Gln Ser Gln Thr Asn
 195 200 205
 Asn Arg Thr Val Asn Thr Leu Ala Val Lys Ser Leu Ala Met Ser Ala
 210 215 220
 Thr Ser Asp Leu Met Ser Asp Thr His Ser Phe Val Arg Leu Asn Asn
 225 230 235 240
 Asn Lys Pro Phe Lys Ile Ser Leu Trp Met Arg Ile Pro Lys Ile Met
 245 250 255
 Lys Ser Asn Thr Tyr Ser Arg Phe Phe Thr Leu Ser Asp Glu Ser Ser
 260 265 270
 Pro Lys Glu Tyr Tyr Ile Ser Ile Gln Cys Leu Pro Asn His Asn Asn
 275 280 285
 Val Glu Thr Val Ile Glu Tyr Asn Phe Asp Gln Ser Asn Leu Phe Leu
 290 295 300
 Asn Gln Leu Leu Leu Ala Val Ile His Lys Ile Glu Met Asn Phe Ser
 305 310 315 320
 Asp Leu Lys Glu Pro Tyr Asn Val Ile His Asp Met Ser Tyr Pro Gln
 325 330 335
 Arg Ile Val His Ser Leu Leu Glu Ile His Thr Glu Leu Ala Gln Thr
 340 345 350
 Val Cys Asp Ser Val Gln Gln Asp Met Ile Val Phe Thr Ile Asn Glu
 355 360 365
 Pro Asp Leu Lys Pro Lys Lys Phe Glu Leu Gly Lys Lys Thr Leu Asn
 370 375 380
 Tyr Ser Glu Asp Gly Tyr Gly Arg Lys Tyr Phe Leu Ser Gln Thr Leu
 385 390 395 400
 Lys Ser Leu Pro Arg Asn Ser Gln Thr Met Ser Tyr Leu Asp Ser Ile
 405 410 415
 Gln Met Pro Asp Trp Lys Phe Asp Tyr Ala Ala Gly Glu Ile Lys Ile
 420 425 430
 Ser Pro Arg Ser Glu Asp Val Leu Lys Ala Ile Ser Lys Leu Asp Leu
 435 440 445
 Asn

45

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1137 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: unknown

50

(ii) MOLECULE TYPE: protein

55

5

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

10

(vi) ORIGINAL SOURCE:

(A) ORGANISM: IMPATIENS NECROTIC SPOT VIRUS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

15

Met Lys Ile Leu Lys Met Cys Glu Leu Leu Val Lys Ile Ser Val Cys
1 5 10 15Thr Leu Val Val Thr Ser Val Ile Leu Ser Phe Met Ala Leu Lys Glu
20 25 30Thr Asp Ala Lys Ile His Val Glu Arg Gly Asp His Pro Glu Ile Tyr
35 40 45

20

Asp Glu Ala Tyr Tyr Asp Arg Ser Val Asp His Lys Asn Glu Ile Leu
50 55 60Asp Thr Leu Ala Glu Met Leu Gln Asn Ala Thr Gly Lys Thr Leu Arg
65 70 75 80

25

Pro Thr Arg Asp Thr Gln Thr Val Leu Ala Asn Asn Glu Val Pro Gln
85 90 95Ser Pro Ser Gly Leu Ser Ser Thr Pro Thr Thr Ile Ser Val Met Asp
100 105 110

30

Leu Pro Asn Pro Cys Leu Asn Ala Ser Ser Leu Thr Cys Ser Ile Lys
115 120 125Gly Val Ser Thr Phe Asn Val Tyr Tyr Gln Val Glu Ser Asn Gly Val
130 135 140

35

Ile Tyr Ser Cys Ile Ser Asp Thr Ile Thr Lys Leu Gly Asn Cys Glu
145 150 155 160Gly Ser Ser Glu Leu Pro Arg Ser Phe Glu Thr Val Pro Val Val Pro
165 170 175Ile Thr Lys Ile Asp Asn Lys Arg Lys Leu Ser Ile Gly Thr Lys Phe
180 185 190

40

Tyr Ile Ile Glu Ser Leu Glu Asn Tyr Asn Tyr Pro Ile Met Tyr Asn
195 200 205Ser Arg Pro Thr Asn Gly Thr Val Ser Leu Gln Ser Val Lys Phe Ser
210 215 220

45

Gly Asp Cys Lys Ile Ser Lys Thr Asn Ile Val Asn Ser Tyr Thr Val
225 230 235 240Ser Leu Thr Thr Pro Glu Lys Ile Met Gly Tyr Val Val Lys Arg Glu
245 250 255

50

Gly Ser Asp Met Ser His Ser Ile Ile Ser Phe Ser Gly Ser Val Ser
260 265 270Leu Thr Phe Thr Glu Glu Asn Met Asp Gly Lys His Asn Leu Leu Cys
275 280 285

55

5

Gly Asp Lys Ser Ser Lys Val Pro Leu Val Asp Lys Arg Val Arg Asp
 290 300
 10 Cys Ile Ile Lys Tyr Ser Lys Asn Ile Tyr Lys Gln Thr Ala Cys Ile
 305 310 315 320
 Asn Phe Ser Trp Phe Arg Leu Ile Met Ile Ala Leu Ile Val Tyr Phe
 325 330 335
 15 Pro Ile Arg Tyr Leu Val Asn Lys Thr Ser Lys Thr Leu Phe Tyr Gly
 340 345 350
 Tyr Asp Leu Leu Gly Leu Ile Thr Tyr Pro Ile Leu Leu Ile Asn
 355 360 365
 20 Tyr Leu Trp Ser Tyr Phe Pro Leu Lys Cys Lys Val Cys Gly Asn Leu
 370 375 380
 Cys Leu Val Thr His Glu Cys Ser Lys Leu Cys Ile Cys Asn Lys Asn
 385 390 395 400
 Lys Ala Ser Glu Glu His Ser Glu Glu Cys Pro Ile Ile Thr Arg Thr
 405 410 415
 25 Ala Glu Lys Asn Lys Lys Tyr Asn Trp Ala Ser Ile Glu Trp Phe His
 420 425 430
 Leu Ile Val Asn Thr Lys Ile Gly Leu Ser Phe Leu Lys Ala Val Thr
 435 440 445
 30 Glu Thr Leu Ile Gly Phe Leu Ile Leu Ser Gln Met Pro Met Ser Met
 450 455 460
 Ala Gln Thr Ala Gln Cys Leu Asp Ser Cys Tyr Tyr Val Pro Gly Cys
 465 470 475 480
 35 Asp Arg Phe Val Thr Asn Arg Tyr Asp Lys Cys Pro Glu Lys Asp Gln
 485 490 495
 Cys Phe Cys Ala Ile Lys Glu Asn Ser Ile Val Glu Ser Asn Phe Leu
 500 505 510
 Thr Asn Val Val Thr Glu Gly Pro Met Asp Cys Ile Pro Tyr Gln Glu
 515 520 525
 40 Cys Lys Gly Arg Ile Thr Glu Asn Ala Leu Val Thr Phe Val Lys Cys
 530 535 540
 Arg Phe Gly Cys Glu Tyr Ala Ser Ile Phe Gln Ser Lys Pro Leu Asp
 545 550 555 560
 45 Asn Gly Phe Leu Glu Tyr Ser Gly Asp Thr Leu Gly Leu Asn Ala Val
 565 570 575
 Asn Leu His Phe Met Lys Arg Leu Arg Asn Gly Ile Ile Asp Phe Tyr
 580 585 590
 50 Asn Lys Thr Glu Lys Tyr Gly Tyr Ile Ser Gly Asp Ala Leu Lys Ser
 595 600 605
 Asn Glu Ser Asp Ile Pro Glu Ser Ile Phe Pro Arg Lys Ser Leu Ile
 610 615 620
 55 Phe Asp Ser Val Ile Asp Gly Lys Tyr Arg Tyr Met Ile Glu Glu Ser

5.

625 630 635 640
 Leu Leu Ser Gly Gly Gly Thr Val Phe Ser Leu Asn Asp Lys Ser Ser
 645 650 655
 10 Ser Thr Ala Gln Lys Phe Val Val Tyr Ile Lys Lys Val Arg Ile Gln
 660 665 670
 Tyr Asp Val Ser Glu Gln Tyr Thr Thr Ala Pro Ile Gln Ser Thr His
 675 680 685
 15 Thr Asp Phe Phe Ser Thr Cys Thr Gly Lys Cys Ser Asp Cys Arg Lys
 690 695 700
 Glu Gln Pro Ile Thr Gly Tyr Gln Asp Phe Cys Ile Thr Pro Thr Ser
 705 710 715 720
 20 Tyr Trp Gly Cys Glu Glu Val Trp Cys Leu Ala Ile Asn Glu Gly Ala
 725 730 735
 Thr Cys Gly Phe Cys Arg Asn Val Tyr Asp Met Asp Gln Ser Phe Arg
 740 745 750
 Ile Tyr Ser Val Ile Lys Ser Thr Ile Lys Ser Glu Val Cys Ile Ser
 755 760 765
 25 Gly Phe Val Gly Ala Lys Cys Phe Thr Val Ser Glu Glu Val Pro Ser
 770 775 780
 Glu Ser Gly Tyr Phe Gln Ala Asp Ile Leu Ala Asp Phe His Asn Asp
 785 790 795 800
 30 Gly Leu Thr Ile Gly Gln Leu Ile Ala His Gly Pro Asp Ser His Val
 805 810 815
 Tyr Ala Gly Asn Ile Ala Arg Leu Asn Asn Pro Ser Lys Met Phe Gly
 820 825 830
 35 His Pro Gln Leu Ser His Gln Gly Asp Pro Ile Phe Ser Lys Lys Thr
 835 840 845
 Leu Asp Thr Asn Asp Leu Ser Trp Asp Cys Ser Ala Ile Gly Lys Lys
 850 855 860
 40 Thr Ile Thr Ile Lys Ser Cys Gly Tyr Asp Thr Tyr Arg Phe Lys Thr
 865 870 875 880
 Gly Leu Asn Gln Ile Ser Asp Ile Pro Val Gln Phe Thr Asp Gln Asn
 885 890 895
 Ser Phe Tyr Met Glu Lys Ile Phe Ser Leu Gly Lys Leu Lys Ile Val
 900 905 910
 45 Leu Asp Leu Pro Ser Glu Leu Phe Lys Thr Val Pro Lys Lys Pro Ile
 915 920 925
 Leu Ser Ser Val Ser Leu Ser Cys Lys Gly Cys Phe Leu Cys Ser Gln
 930 935 940
 50 Gly Leu Arg Cys Ala Ala Ser Phe Ile Ser Asp Ile Thr Phe Ser Ala
 945 950 955 960
 Arg Leu Thr Met Lys Gln Cys Ser Leu Ser Thr Tyr Gln Ile Ala Val
 965 970 975

55

5	Lys Lys Gly Ala Asn Lys Tyr Asn Leu Thr Met Phe Cys Thr Ser Asn 980 985
	Pro Glu Lys Gln Lys Met Ile Ile Glu Pro Glu Gly Asp Lys Ser Tyr 995 1000 1005
10	Ser Val Glu Ala Leu Val Asp Ser Val Ala Val Leu Glu Pro Glu Asn 1010 1015 1020
	Ile Ile Asp Gln Asn Asp Gln His Ala His Glu Glu Gln Gln Tyr Asn 1025 1030 1035 1040
15	Ser Asp Thr Ser Val Trp Ser Phe Trp Asp Tyr Val Lys Ser Pro Phe 1045 1050 1055
	Asn Phe Ile Ala Ser His Phe Gly Ser Phe Phe Asp Thr Val Arg Val 1060 1065 1070
	Val Leu Leu Ile Leu Phe Val Phe Ala Leu Ala Tyr Leu Cys Ser Ile 1075 1080 1085
20	Val Ala Thr Met Cys Arg Gly Tyr Val Arg Asn Lys Ser Tyr Lys Thr 1090 1095 1100
	Lys Tyr Ile Glu Asp Thr Asn Asp Tyr Ser Leu Val Ser Thr Ser Ser 1105 1110 1115 1120
25	Gly Lys Asp Thr Ile Thr Arg Arg Arg Pro Pro Leu Asp Phe Ser Gly 1125 1130 1135
	Ile

Claims

1. Recombinant INSV DNA constructs comprising a DNA sequence coding for transcription into
 - a) an RNA sequence of an INSV or an RNA sequence homologous thereto ;
 - b) an RNA sequence of an INSV or an RNA sequence homologous thereto capable of encoding for an INSV protein or a part thereof , in which one or more codons have been replaced by synonyms , or an RNA sequence homologous thereto ; or
 - c) an RNA sequence complementary to an RNA sequence according to a) or b),which INSV DNA is under-expression control of a promoter and a terminator capable of functioning in plants.
2. A DNA construct according to Claim 1, wherein the INSV DNA sequences code for transcription into:
 - i) the viral S RNA nucleotide sequence from 1 to 3017 (SEQ. ID No.1)
 - ii) the viral S RNA nucleotide sequence from position 25 to 3017 (SEQ. ID No.2);
 - iii) the viral S RNA nucleotide sequence from 87 to 1436 (SEQ. ID No.3);
 - iv) the viral S RNA nucleotide sequence from 2080 to 2868 (SEQ. ID No.4);
 - v) the viral S RNA " pan-handle " structure comprising :
 - a) a first nucleotide sequence of from about 30 to about 36 nucleotides in length from the 5' end of the viral S RNA
 - and
 - b) a second nucleotide sequence of from about 30 to about 36 nucleotides in length from the 3' end of the viral S RNA
 - vi) the viral S RNA nucleotide sequence from 1437 to 2079; (SEQ ID No. 7)
 - vii) the viral S RNA nucleotide sequence from 1440 to 2041; (SEQ ID No.8)
 - viii) the viral complementary S RNA nucleotide sequence from 1 to about 3017; (SEQ ID No.9)
 - ix) the viral complementary S RNA nucleotide sequence from 1 to 2993; (SEQ ID No.10)
 - x) the viral complementary S RNA nucleotide sequence from 150 to 938; (SEQ ID No.11)
 - xi) the S RNA nucleotide sequence from 1581 to 2930 of the viral complementary S RNA strand; (SEQ ID

No.12);

xii) the viral complementary S RNA secondary structure having a nucleotide sequence of 642 nucleotides from 939 to 1580; (SEQ ID No.13)

xiii) S RNA nucleotide sequence from 87 to 1436 in which one or more codons have been replaced by their synonyms;

xiv) S RNA nucleotide sequence from 2080 to 2868 in which one or more codons have been replaced by their synonyms;

xv) the M RNA nucleotide sequence from 1 to 4970 (SEQ ID No.14);

xvi) the M RNA sequence from 86 to 997 (SEQ ID No.15);

xvii) the M RNA sequence of the intergenic region from 998 to 1470 (SEQ ID No.16);

xviii) the M RNA sequence from 1471 to 4884; (SEQ ID No. 17)

xix) the M RNA "pan-handle" structure comprising : a) a first nucleotide sequence of from about 30 to about 36 nucleotides in length from the 5' end of the viral M RNA

and

b) a second nucleotide sequence of from about 30 to about 36 nucleotides in length from the 3' end of the viral M RNA;

xx) the complementary viral M RNA sequence from 1 to 4970; (SEQ ID No.20)

xxi) the complementary viral M RNA sequence from position 87 to position 3500 of the complementary viral M RNA sequence; (SEQ ID No.21)

xxii) the complementary viral M RNA sequence from position 3974 to 4885 (SEQ ID No.22)

xxiii) RNA sequences homologous to the nucleotide sequences defined under i) to xii) and xv) to xxii) hereinabove.

xxiv) fragments of sequences defined under i) to xxii) hereinabove.

3. A DNA construct according to Claim 1, wherein the DNA sequence codes for transcription into INSV-RNA sequences of a pan-handle, or into RNA sequences homologous thereto.

4. A DNA construct according to Claim 1 wherein the DNA sequence codes for transcription into INSV-RNA sequences of a pan-handle wherein the pan-handle structure comprises two complementary strands comprising 36 nucleotides in length.

5. A DNA construct according to Claim 1, wherein the DNA sequence codes for transcription into INSV RNA sequences of an open reading frame in viral complementary sense, or into corresponding RNA sequences in which one or more codons have been replaced by synonyms thereof, or into RNA sequences homologous thereto.

6. A DNA construct according to Claim 1, wherein the the DNA sequence codes for transcription into INSV-RNA sequences of a secondary structure, or into RNA sequences homologous thereto.

7. A DNA construct according to Claim 1 wherein the DNA sequence codes for transcription into INSV-RNA sequences, or into INSV-RNA sequences in which one or more codons have been replaced by synonyms thereof, or into RNA sequences homologous thereto of at least 15 nucleotides.

8. A DNA construct according to Claim 6, wherein the DNA sequence codes for transcription into INSV-RNA sequences in which one or more codons have been replaced by synonyms thereof, or into RNA sequences homologous thereto of at least 400 nucleotides.

9. A DNA construct according to Claim 1, wherein the DNA sequence codes for transcription into a combination of the 5' and 3' terminal sequences of the viral S, M or L RNA respectively.

10. A DNA construct according to Claim 1, wherein the promoter is a viral, fungal, bacterial, animal or plant-derived promoter capable of functioning in plant cells.

11 A DNA construct according to Claim 9, wherein the terminator is a viral fungal, bacterial, animal or plant-derived terminator capable of functioning in plant cells.

12 A plant comprising in its genome a DNA construct in accordance with Claim 1.

13 A probe comprising a single or double stranded oligonucleotide sequence complementary to an RNA sequence of an INSV.

14 A probe according to Claim 13, wherein the oligonucleotides are complementary to an INSV S RNA sequence, an INSV M RNA nucleotide sequence or to fragments of such sequences comprising at least 12 nucleotides.

15 A probe according to Claim 12 or Claim 13, wherein the oligonucleotide sequence has from 12 to 800 nucleotides.

16 A process of preparing plants according to Claim 11, which comprises

a) inserting into the genome of a plant cell a DNA construct of Claim 1,

b) obtaining transformed cells;

c) regeneration from the transformed cells genetically transform d plants.

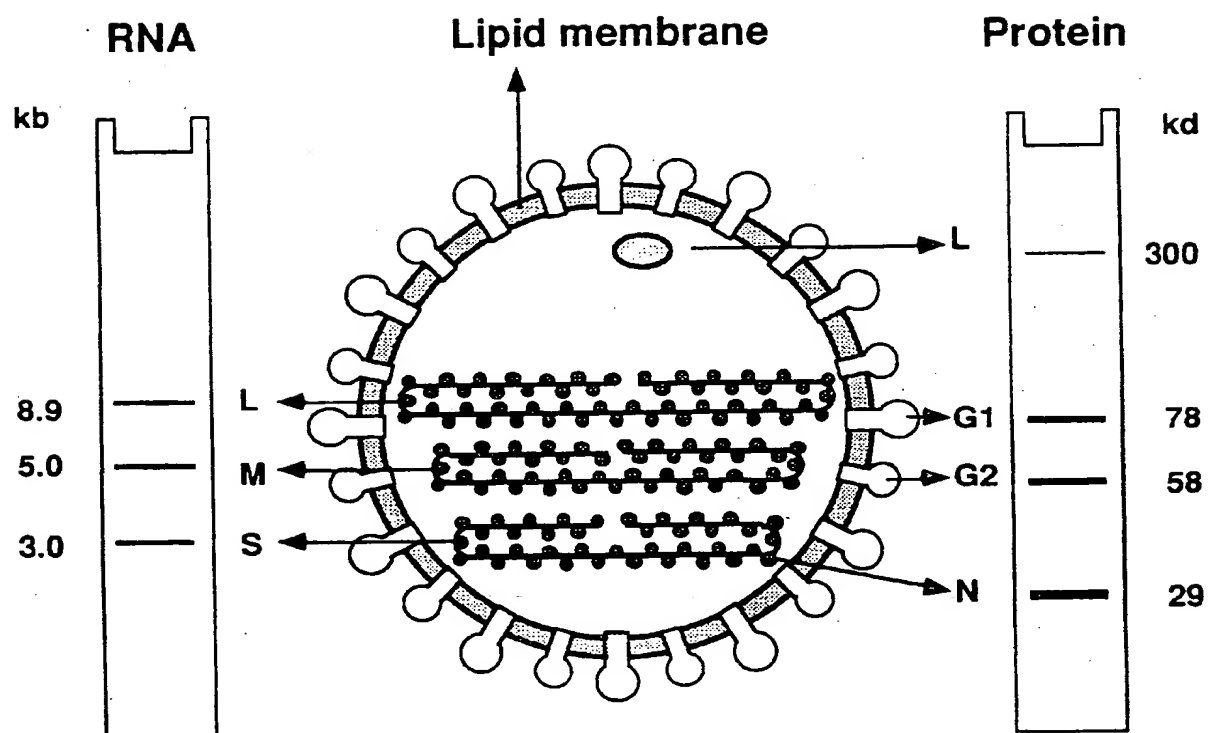


Figure 1

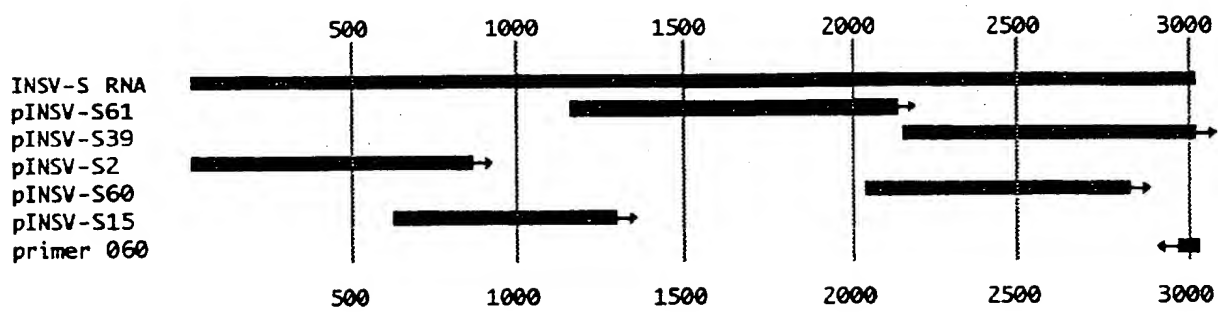


Figure 2

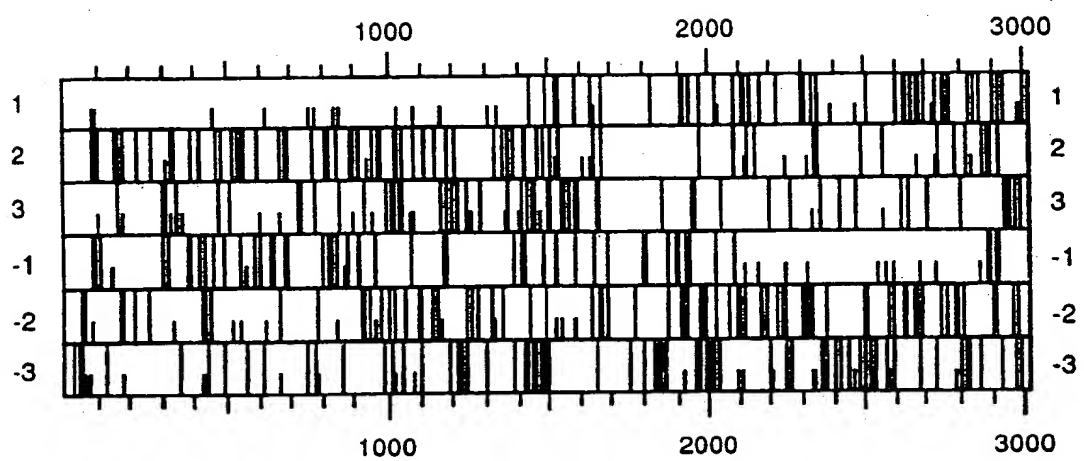


Figure 3

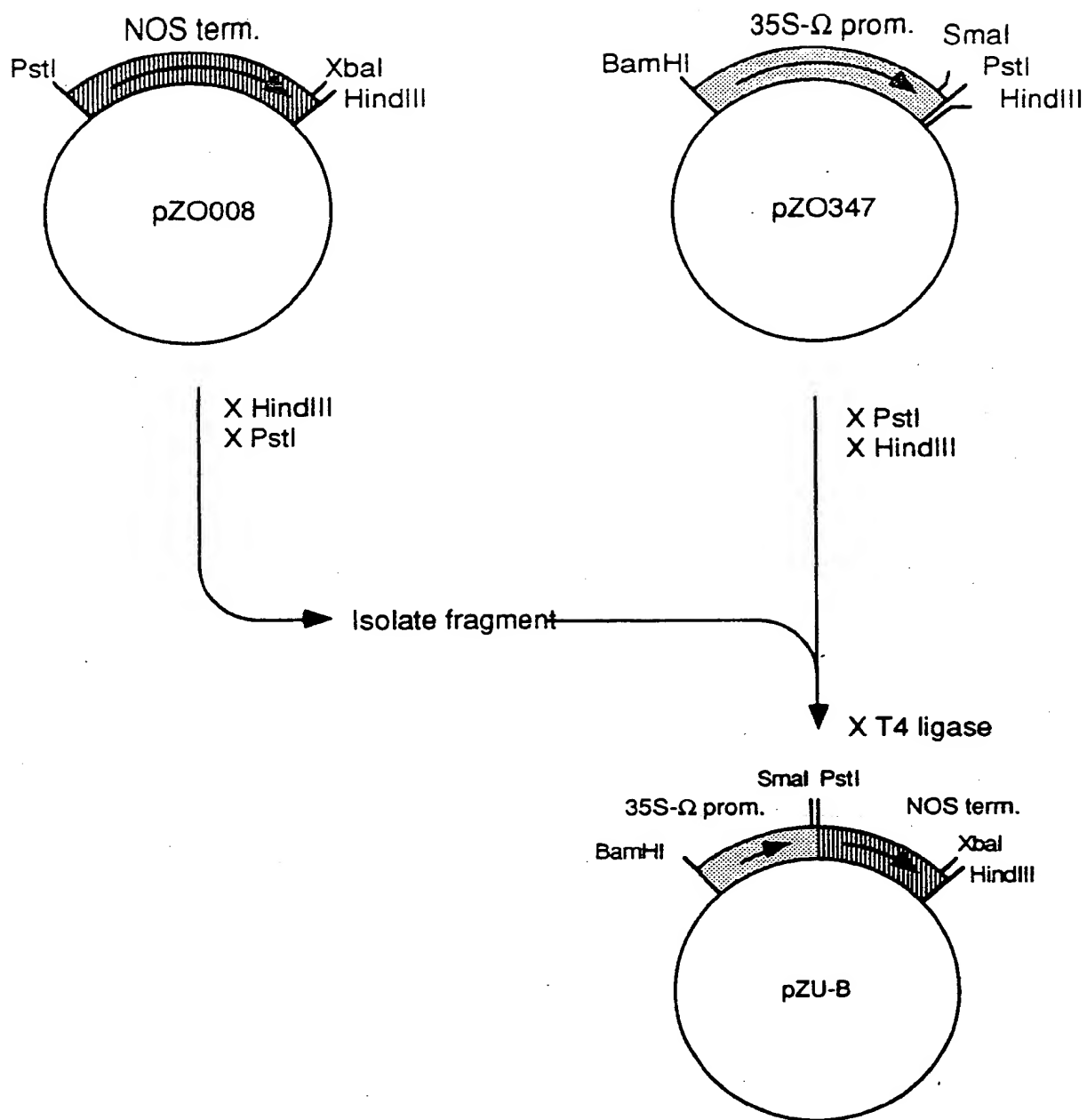


Figure 4

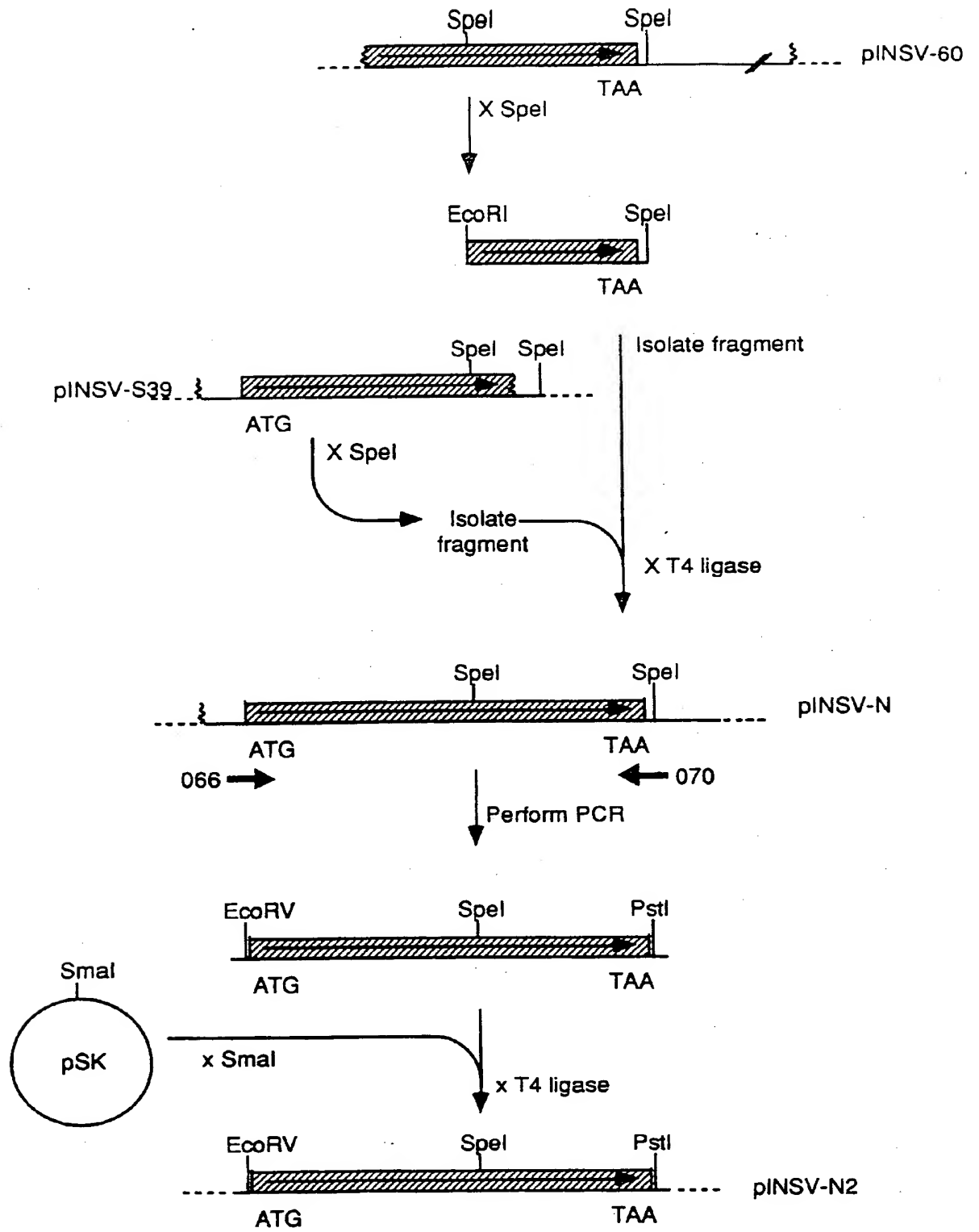


Figure 5

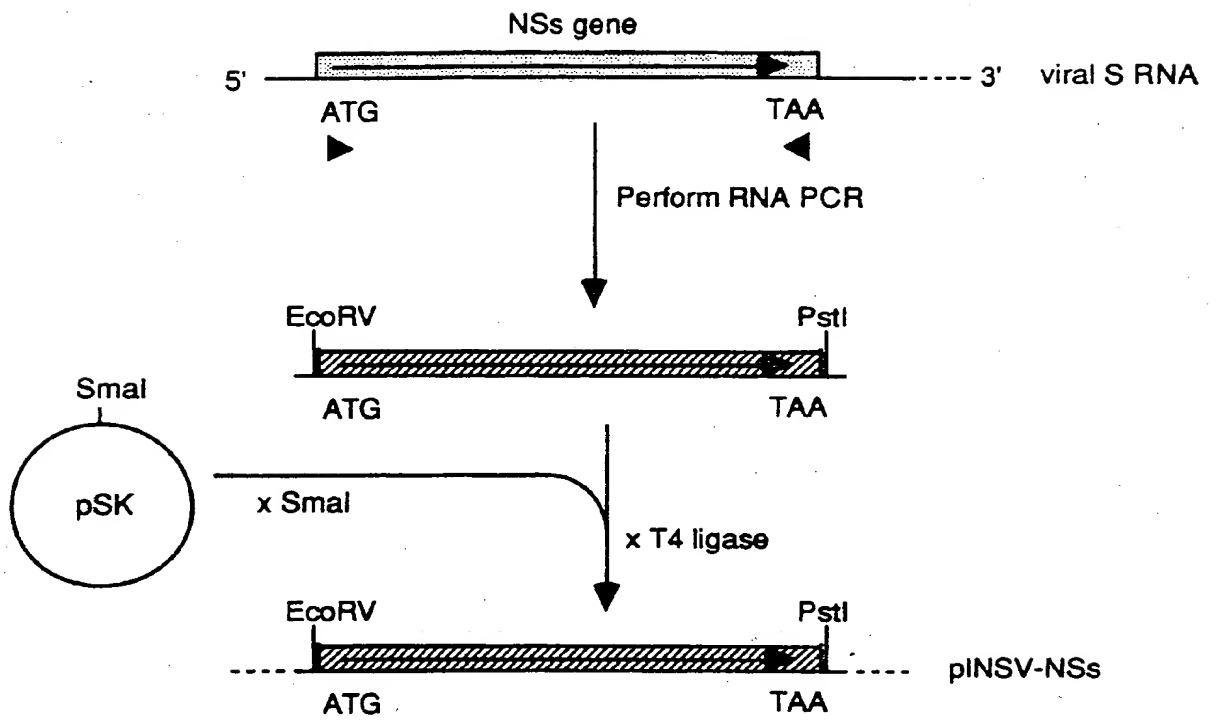


Figure 6

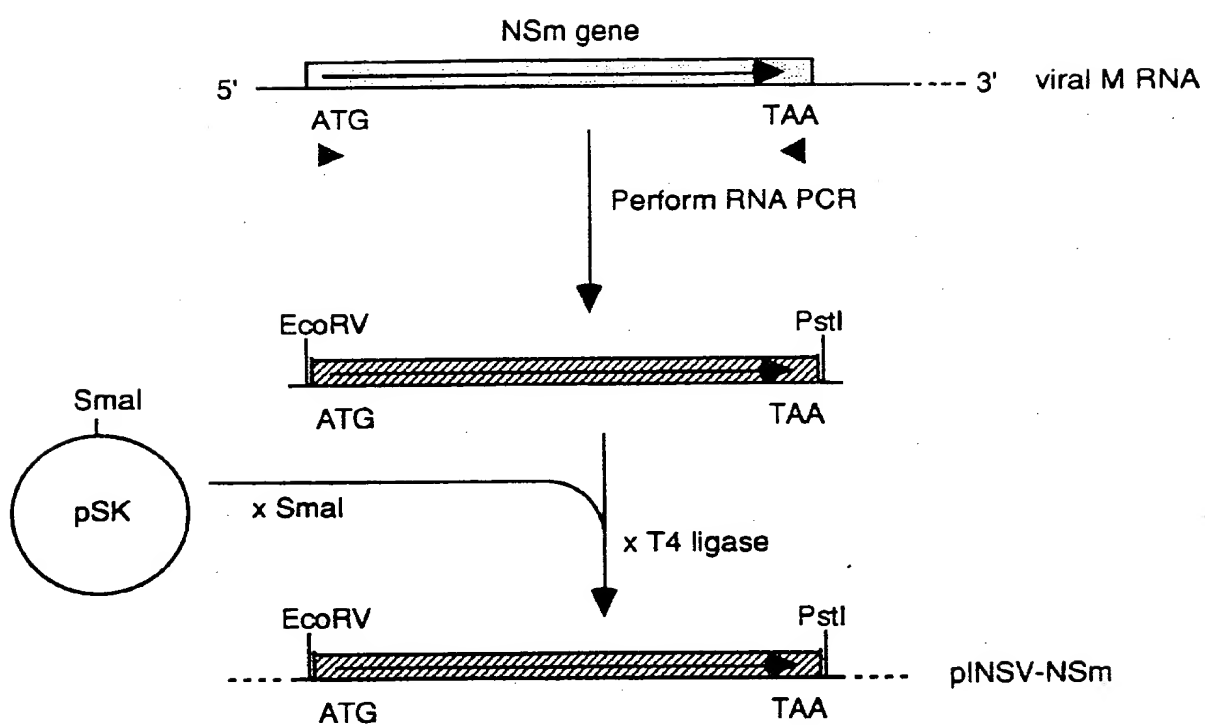


Figure 7

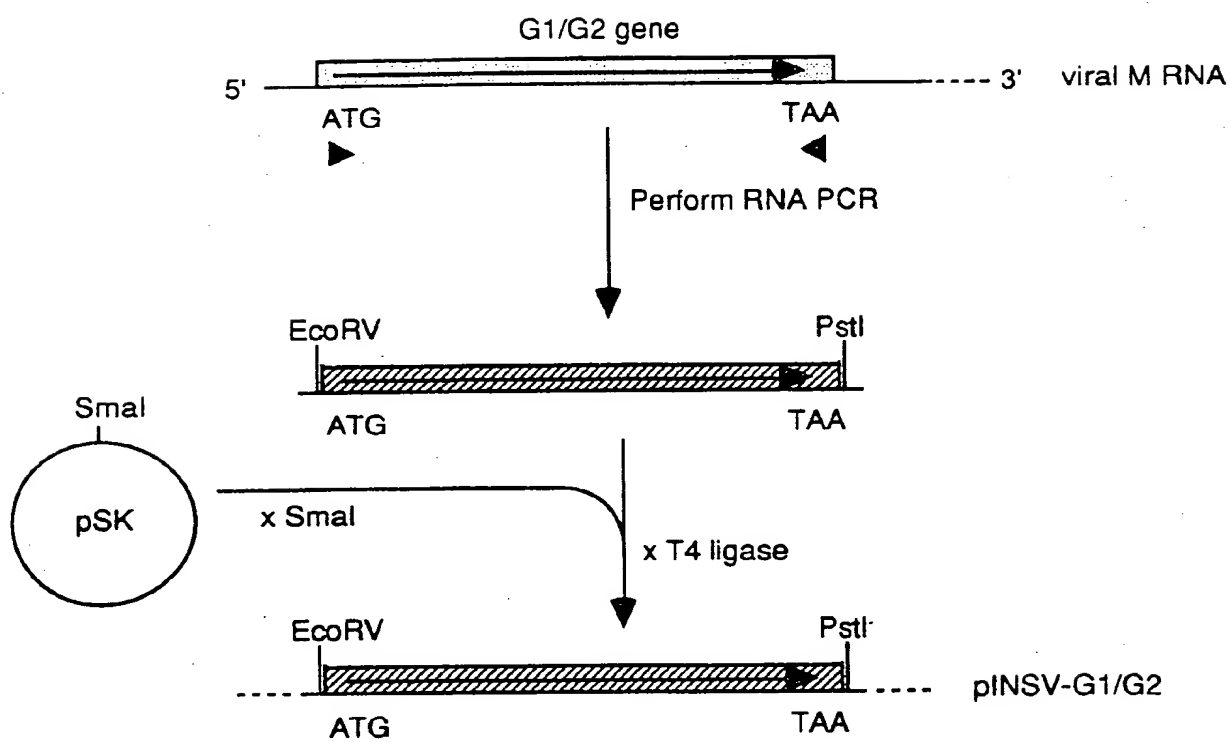


Figure 8

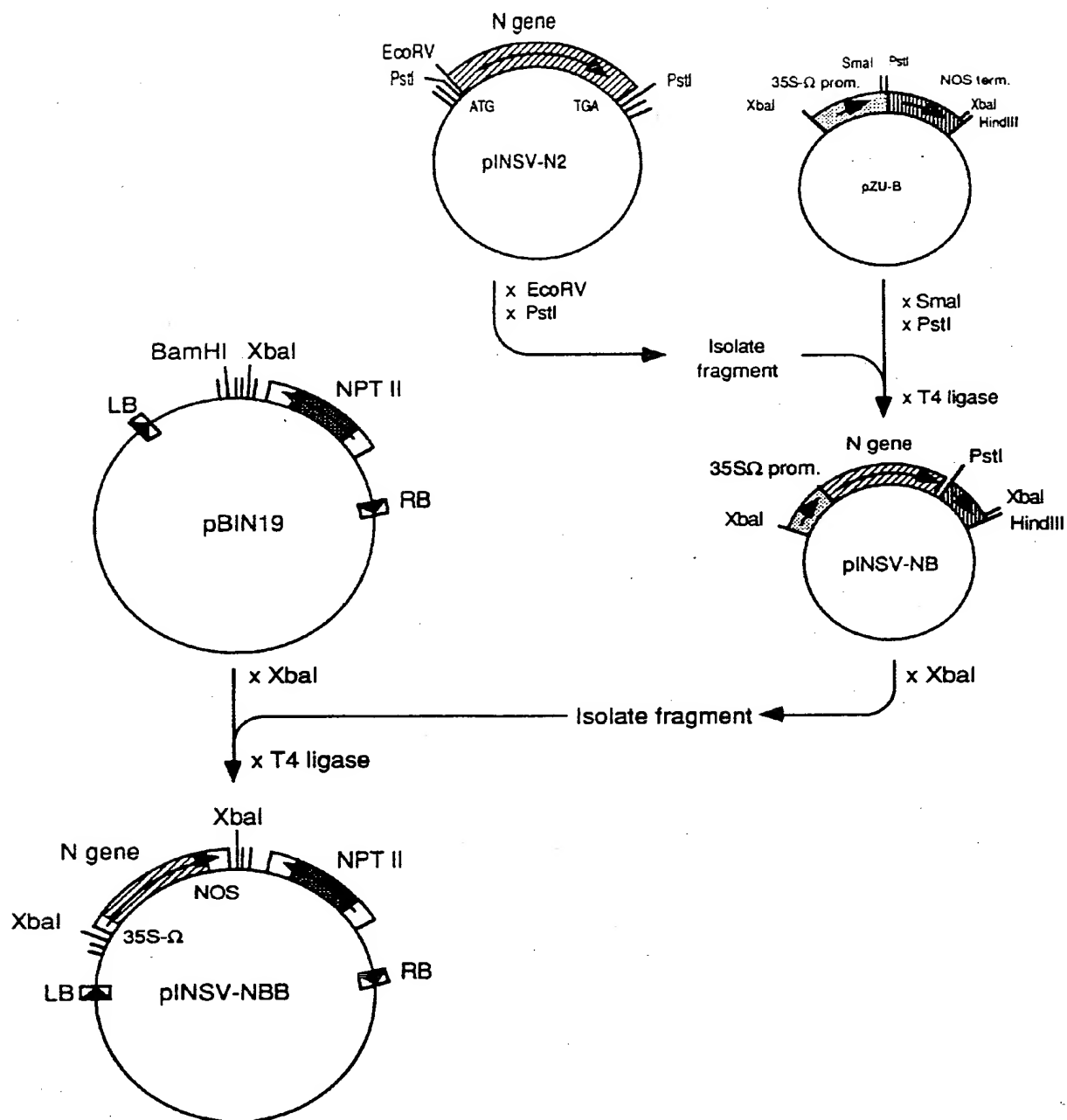


Figure 9

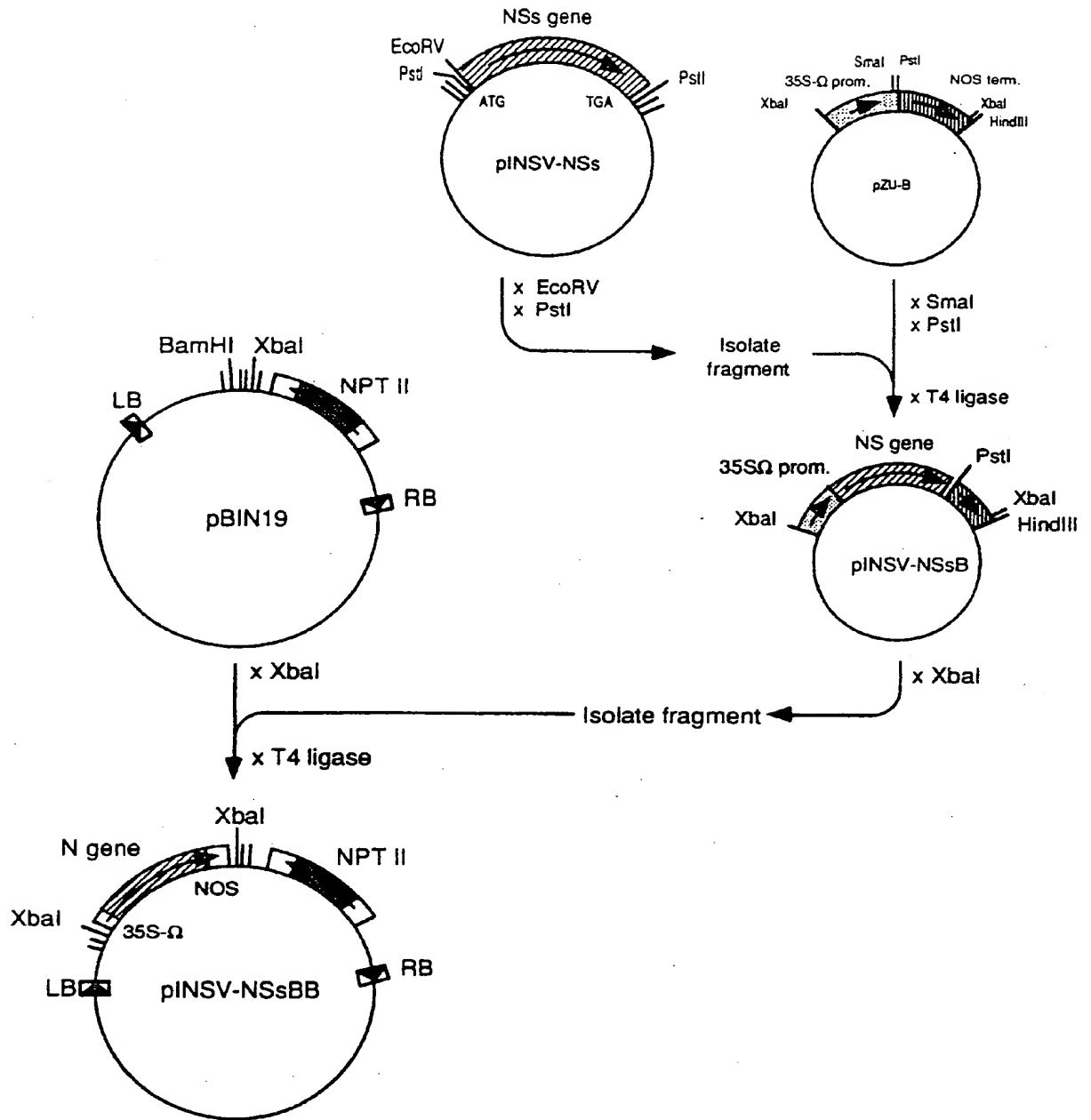


Figure 10

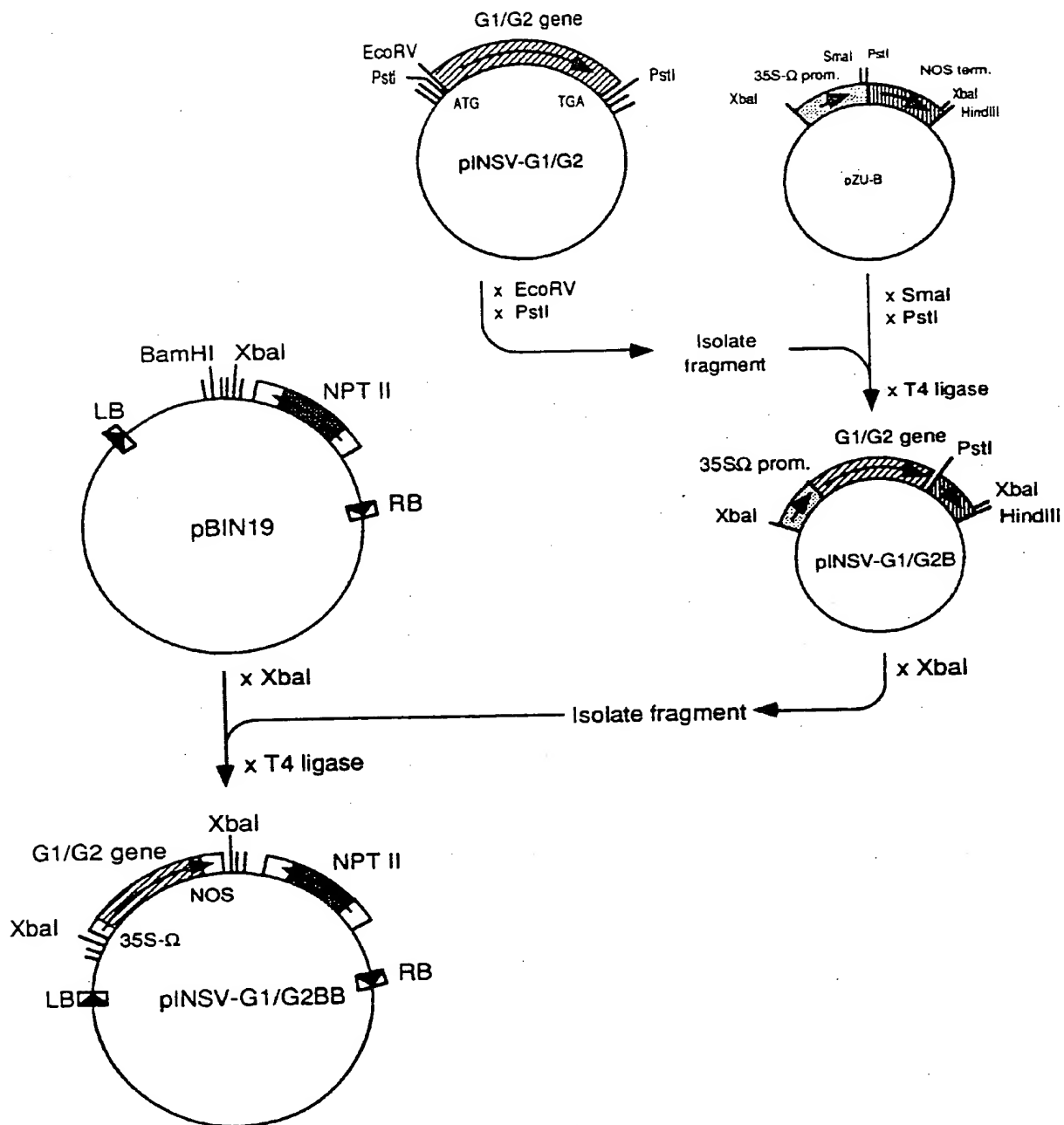


Figure 11

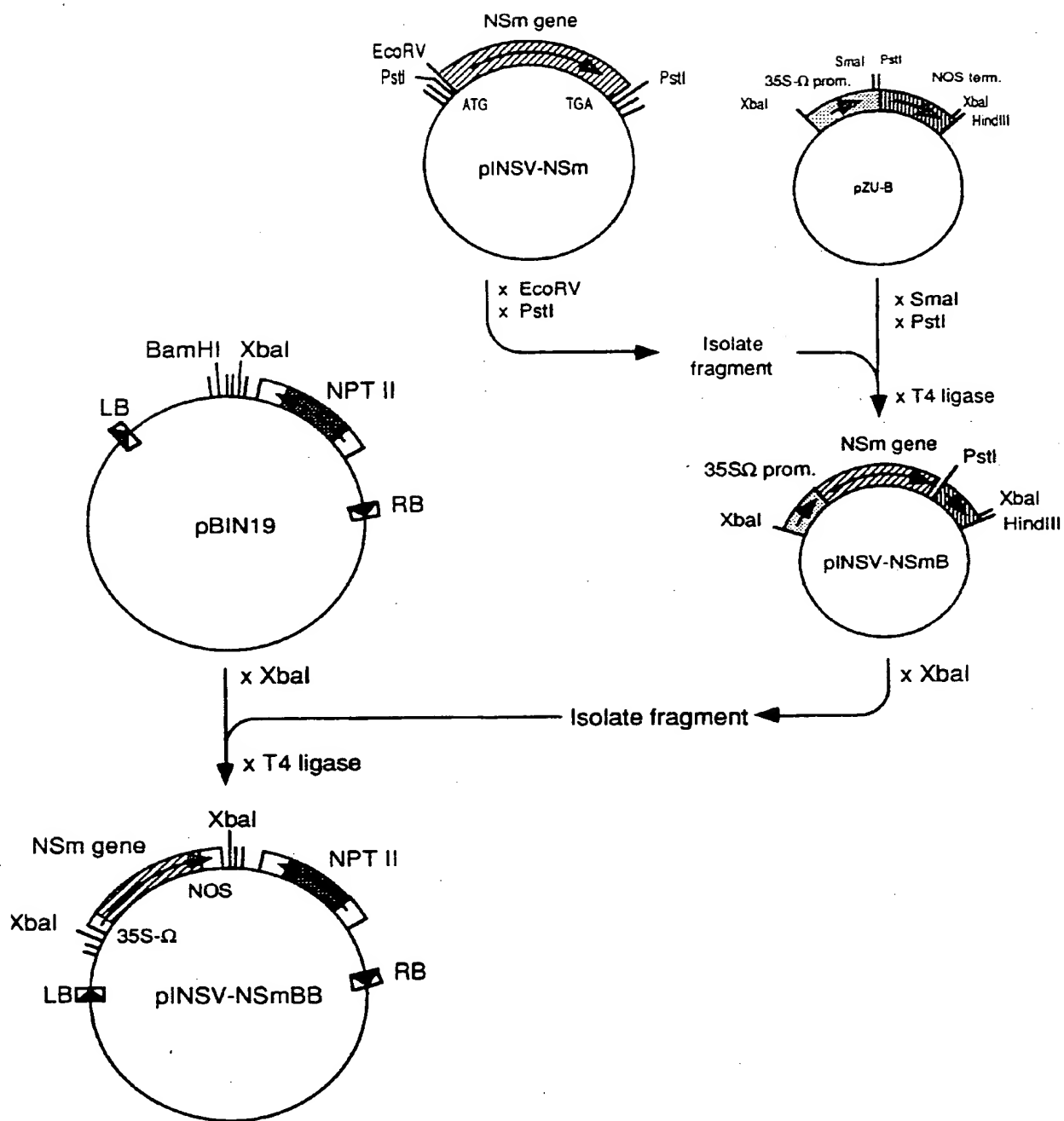


FIGURE 12

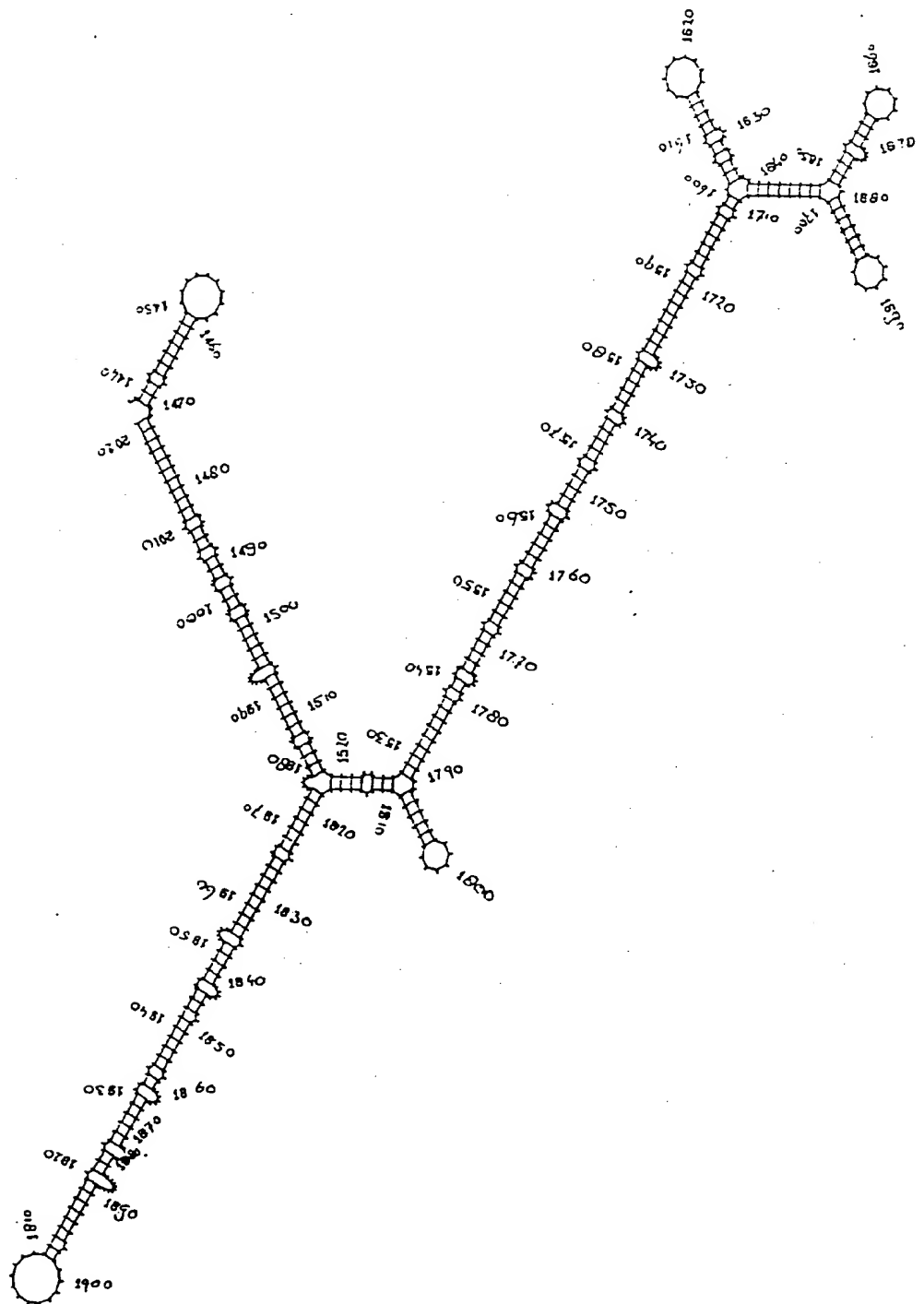


Figure 13



C12N15/82

(11) Publication number: 0 566 525 A3

(12)

EUROPEAN PATENT APPLICATION

(21) Application number: 93810190.4

(51) Int. Cl.⁵: C12N 15/40, C12N 15/82,
C12Q 1/70, A01H 5/00

(22) Date of filing: 16.03.93

(30) Priority: 19.03.92 GB 9206016

(43) Date of publication of application:
20.10.93 Bulletin 93/42

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE IT LI LU NL
PT SE

(88) Date of deferred publication of search report:
08.12.93 Bulletin 93/49

(71) Applicant: SANDOZ LTD.
Lichtstrasse 35
CH-4002 Basel (CH)
(84) BE CH DK ES FR GB GR IE IT LI LU NL PT SE

(71) Applicant: SANDOZ-PATENT-GMBH
Humboldtstrasse 3
D-79539 Lörrach (DE)
(84) DE

(71) Applicant: SANDOZ-ERFINDUNGEN
Verwaltungsgesellschaft m.b.H.
Brunner Strasse 59
A-1230 Wien (AT)
(84) AT

(72) Inventor: Van Grinsven, Martinus Quirinus
Joseph Marie
Wezenland 5
NL-1602 MA Enkhuizen (NL)
Inventor: De Haan, Petrus Theodorus
Kruideel 34
NL-1602 GL Enkhuizen (NL)
Inventor: Gielen, Johannes Jacobus
Ludgerus
Jan Gooskaai 73
NL-1602 GC Enkhuizen (NL)
Inventor: Peters, Dirk
Edelmanlaan 4
NL-6703 EX Wageningen (NL)
Inventor: Goldbach, Robert Willen
Hollandseweg 159
NL-4705 BC Wageningen (NL)

(54) Recombinant tospovirus DNA constructs and plants comprising such constructs.

(57) Recombinant Impatiens Necrotic Spot Virus (INSV) DNA constructs comprising an INSV DNA coding for transcription into INSV RNA sequences or into RNA sequences related thereto, the use of such DNA constructs to transform plants having reduced susceptibility to INSV infection and probes for the isolation of INSV or diagnosis of plant INSV related diseases.

EP 0 566 525 A3



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP 93 81 0190
Page 2

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
A	ANNUAL REVIEW OF PHYTOPATHOLOGY vol. 30, 1992, pages 315 - 348 GERMAN, T.L., ET AL. 'Tospovirus diagnosis molecular biology phylogeny and vector relationships' * page 317 - page 319 * ---	13-15	
A	CHEMICAL ABSTRACTS, vol. 117, 1992, Columbus, Ohio, US; abstract no. 185748, LAW, M.D., ET AL. 'Nucleotide sequence of the 3' non-coding region and N gene of the S RNA of a serologically distinct tospovirus' * abstract * & J. GEN. VIROL. vol. 72, no. 10, 1991, pages 2597 - 2601 ----	13-15	
A	J. GEN. VIROL. vol. 71, 1990, pages 933 - 938 LAW, M.D., ET AL. 'A tomato spotted wilt virus with a serologically distinct N protein' * the whole document * ----- -/--	1-16	
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 24 SEPTEMBER 1993	Examiner MADDOX A.D.
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons</p> <p>* : member of the same patent family, corresponding document</p>			

EPO FORM 1503 (11.92) (P0601)

